THE APPROVAL MOTIVE

Studies in Evaluative Dependence

THE APPROVAL MOTIVE

Studies in Evaluative Dependence

Douglas P. Crowne, UNIVERSITY OF CONNECTICUT

David Marlowe, HARVARD UNIVERSITY

John Wiley & Sons, Inc., New York · London · Sydney

Library of Congress Catalog Card Number 64–25893

Printed in the United States of America

TO THE MEMORY OF

Shephard Liverant and Alvin Scodel

Friends, colleagues, and teachers

Preface

Not long after personality tests were introduced, psychologists discovered that they were exasperatingly unsuccessful What had seemed at first like such a simple task—predicting a person's behavior from his responses to questions about his thoughts, feelings, and symptoms—turned rapidly into a miasma of complexity, and the goal became very nearly lost to view Altogether too many people just didn't behave as their test responses said they should, and to the psychologists of the 1920's it must have seemed that their subjects were miserably uncooperative It took some time for the early enthusiasm for tests to lose its luster In the meantime, the situation for the psychologist resembled nothing so much as an armed encounter—the subject resistively glowering across the psychologist's desk, the psychologist struggling to maintain his air of professional imperturbability and not being quite able to suppress a look of puzzled astonishment at such cavalier disdain for the wonders of his marvelous brainchild, the personality inventory It is hardly surprising that the brilliance of the vision would for a time obscure the subject's stake in self-disclosure and the uses to which it would be put

Grudgingly, psychologists acknowledged that a problem existed, and it was first laid at the subject's doorstep But passing the blame was of little help, and entreaties were of no avail If personality tests were to survive, the next step was clear the

subject's lack of frankness would have to be revealed and his responses pared down to the honest bone

With increasing sophistication, psychologists attempted to rid tests of dissimulation by identifying and correcting for censorship and distortion, and from the mid-1930's on a number of procedures were developed to accomplish this purpose A host of questions emerged in the process Did subjects "fake good" or "bad" according to their purposes? Were test responses distorted consciously or unconsciously? How much of a role did carelessness play? Was defensiveness a factor? Could it be that subjects responded to tests with a "set" or "habit" of acquiescence or denial? Or, perhaps, did some subjects depict themselves in a socially acceptable manner? Several devices in the form of special keys for scoring personality tests and special scales to detect dissimulation were created, they met with only moderate success The problem lingered on, and psychologists now began to take a more serious look at the questions themselves The direct assault on the problem was called back, that concentration might be focused on the most basic question of all What are the sources of variance in personality test responses?

In casting about for the answer to this question, the concept of response set was introduced, embodying the notion that subjects bring to testing sets, or habitual response preferences, such as the tendency to indiscriminantly agree with test items If response sets like this were present in any significant degree, they would of course hopelessly obscure the subject's responses to the content of the items themselves Thus, the investigation of the preeminent set, acquiescence—its properties and the extent of its influence on personality test responses—blossomed Following close on the heels of acquiescence was the concept of social desirability, referring ambivalently to personality test *items* as socially desirable or undesirable statements about oneself and to the tendency of *subjects* to endorse socially desirable options Along with these developments came the beginnings of speculation about *why*—about the meaning of dissimulation, or acquiescence, or social desirability

Five years ago, our curiosity kindled by a stubborn problem and some perspicacious recent developments, we set out to explore the social-desirability response set As a descriptive ac-

count of test response distortion it had an intuitive plausibility about it, and it fitted rather neatly into an explanation which we thought bore some promise That explanation was simply this people describe themselves in favorable, socially desirable terms in order to achieve the approval of others

This book describes the research that followed In a series of studies beginning in the summer of 1959 we developed a social-desirability questionnaire and attempted first to test the implications of the concept of need for approval for personality testing To do so we had to establish that the need for approval was a valid inference from socially desirable test responses, we had to demonstrate that a presumed measure of the approval motive would predict approval-seeking behaviors in situations other than testing We sought, then, to establish that behavior which had long seemed a nuisance is meaningful, goal-directed, and related to the individual's behavior in other situations

In the process, we moved a long way from the domain of personality tests to seek the predictive utility of the approval motive in experimental situations in the laboratory and in life situations outside

It is the purpose of *The Approval Motive* to bring these experiments together, describe them, and interpret them The book is divided into four parts Part I takes up the problem of the social-desirability response set and our initial efforts to measure and explain it In Parts II and III are the experiments designed to test the implications of the approval motive—studies of compliance, influencibility, and conformity, and some explorations in the realm of personal vulnerability and defensiveness Finally, Part IV attempts a theoretical integration of the findings, bringing the approval motive within the fold of a more inclusive theory of personality

Many of the major experiments described in this book—approximately 60 per cent of them—have been previously published, they have been reported in various psychological journals as they were completed They are brought together here with several new studies so that we might report an integrated body of findings and develop more completely than one can ever do in a series of articles the theory devised to explain them The approval motive is not validated by the results of a single experi-

ment, its claim to validity rests on a network of related findings To present that network and the way it evolved is the goal of *The Approval Motive*

This has been in large measure a collaborative enterprise with students and colleagues We take great pleasure in recording our indebtedness to them Their ideas, critical comments, enthusiasm, and most of all, their research contributions have made this book possible The findings we report are theirs just as much as ours Our thanks go to Stephen P Bank, Dr Christopher E Barthel, Russell S Beecher, Jonathan B Cook, Dr Lane K Conn, Martin Davis, Anthony N Doob, Dr John M Rosenfeld, Arthur R Salman, Lawrence Stifler, Dr Bonnie R Strickland, and numerous other members of research teams and informal research seminars In addition, we would like to express our appreciation to Dr Thomas A Tutko for permitting us to describe his doctoral dissertation in Chapter 11 To his adviser, Dr Lee B Sechrest, also go our thanks for sharing with us his perceptive ideas on the response-set problem and criticizing early versions of the manuscript.

Various chapters were discussed with Drs David L Horton, Robert Rosenthal, Julian B Rotter, and Mark W Stephens, the latter two read the entire manuscript We earnestly hope that the book reflects in some small way how much we profited from the help of these colleagues

Two summers of research and writing were generously supported by research grant G25113 from the National Science Foundation The major theoretical chapters and several others were written in happy freedom from the usual cares of the academic year Some of the experiments were also supported by the grant, we are pleased to acknowledge this assistance

The book is dedicated to the memory of our friends, colleagues, and teachers, Drs Shephard Liverant and Alvin Scodel We are deeply grateful for the wisdom and friendship they gave us Finally, to our wives, Susan and Lynn, goes our warm appreciation for generous help and gentle forbearance

Evanston, Illinois
Cambridge, Massachusetts
August 1964

Douglas P Crowne
David Marlowe

Contents

THE APPROVAL MOTIVE

Studies in Evaluative Dependence

part I The problem of the social-desirability response set

1

Response set and personality assessment: some basic issues

ORIGINS

It is an observation of very long standing that scores on personality questionnaires are influenced by factors other than the manifest content of the items Simply from the fact that an individual responds *false* to the item, "Once in a while I think of things too bad to talk about," it cannot be determined whether his thoughts are indeed pure, whether he has bad thoughts but refuses for any of a number of reasons to divulge that fact, or whether he is perhaps generally disposed to disagree with most personality-test items irrespective of their content As a matter of fact, even a *true* response to such an item would leave the possibility that the respondent was displaying a tendency to agree with most statements regardless of their content

The omnipresence of this problem—that responses to personality-test items cannot always be taken at face value—has been a perennial source of frustration to psychologists in their quest for the grail of successful prediction Except for a few dissenters (Allport, 1928, Cady, 1923, Rosenzweig, 1934) a traditional staple of the early test constructors' stock-in-trade was the belief that the *content* of personality-test items is the major, if not the sole, determinant of responses Responses to test items were viewed as an effective substitute for the direct obser-

vation of actual behavior Thus, it was assumed that the subject *could* report accurately how he typically behaved and, further, that he *would* willingly reveal his behavioral dispositions in his answers to test items The obvious strategy in constructing personality questionnaires was to ask direct and undisguised questions about typical behavior, characteristics, and feelings—questions to which the respondent would then commit himself with "yes" or "no" (e g , "Do you frequently have headaches?" "Are you easily disappointed?" "Are you afraid of high places?") This was the procedure followed in many of the early tests Of course, if test responses were really isomorphic to behavior, then it would little matter if the content of items, or indeed of the entire test, were transparent to the subject or required him to admit to socially undesirable traits or symptoms

As short-cut devices and substitutes for behavioral observation, personality tests proved, disappointingly, to have little predictive validity (Ellis, 1946) Rather than abandoning the infant testing movement, however, psychologists turned a mistrustful eye on its conceptual parentage—the presuppositions on which personality tests were based Doubt was cast on whether subjects *would* reveal their traits, attributes, and characteristic ways of responding "How honest is the subject?" became the open question More, though, than test assumptions was skeptically regarded an invidious reflection was obviously cast on the subject

Understandably, from the point of view of the psychologist, this new and jaundiced view of the testee's motives was accompanied by the parallel development of the idea—tenaciously held for many years—that tendencies to censor or dissimulate constitute a source of *error,* rather than valid, variance A number of strategems were then employed to rid tests of such a pernicious influence Exhortations to honesty in instructions to the subject, assurances of anonymity (violated, of course, by the examiner), and sundry attempts to detect inconsistency (Cady, 1923) or improbable asseverations of desirable attributes (Hartshorne & May, 1928, Meehl & Hathaway, 1946) were variously undertaken to reveal, if not to overcome, the failure

of the subject to produce the frank and uncensored test responses desired by the psychologist

THE UNTHEORY OF RESPONSE DISSIMULATION
EMPIRICAL APPROACHES

It seems to have been conceded during the 1930's that subjects could not be successfully prevailed upon, or trapped into, revealing themselves on self-report instruments, and research activity turned to efforts to identify dissimulators and to correct for pre-existing tendencies to conceal The work of Humm and his associates (Humm & Humm, 1944, Humm, Storment, & Iorns, 1939, Humm & Wadsworth, 1935) is a classic example In the industrial screening of job candidates, both the extremes of personally enhancing dissimulation and unrealistic self-criticism were held to lead to reduced predictive validity This followed, of course, from the assumption that item content mediates the relation between test responses and "real life" behavior Humm and his colleagues accordingly devised a scheme to identify such subject tendencies on the Humm-Wadsworth Temperament scale What they came up with was the "no count" index, which basically involved a comparison of denial responses with an empirical criterion—the base rate of *no* (or *false*) responses in the population

The development of the most sophisticated of personality inventories, the Minnesota Multiphasic Personality Inventory, capitalized on the growing awareness of the response-dissimulation problem and the efforts to detect deception represented in the works of Humm et al and Hartshorne and May Ruch (1942), also, was a forerunner in recognizing the problem It was Ruch who first developed a "fake" key by contrasting the responses of subjects instructed to fake with those of subjects in a standard instructional condition on the Bernreuter Personality Inventory

In the construction of the MMPI, three special validity scales (*F*, *L*, and *K*) were incorporated to detect and, in the case of the *K* scale, to suppress, defensive or "plus-getting" tendencies—or, in more extreme language, dispositions to "fake good" or to

"fake bad" The *K* scale (McKinley, Hathaway, & Meehl, 1948, Meehl & Hathaway, 1946), which attempted to correct for these undesirable dispositions to conceal or to dissimulate, represents a culmination of these efforts

We may note here that in the development of the MMPI the position was taken (*cf* Meehl, 1945) that personality-test items need not have the same meaning for all subjects and that particular responses may frequently be an inaccurate index of what a person is really like Meehl argued that test responses are significant instances of verbal behavior in their own right regardless of their correspondence, or lack of it, with behavior Approaching the problem this way, a personality scale is constructed by selecting items on empirical grounds with the single requirement that they discriminate reliably between criterion groups, irrespective of the meaning a subject "projects" into a given item So long as the use of the MMPI was restricted to the assignment of individuals to diagnostic categories on strictly empirical grounds, the problem of the sources of response variation was of secondary importance Indeed, no problem could be said to exist, the issue was simply side-tracked, with ample justification With this perspective on personality-test construction as a point of departure, and given that prediction, empirical or otherwise, is imperfect, the MMPI validity scales served to identify records that were atypical because the subject was careless, failed to understand the instructions, lied, or attempted to present an unrealistically good or bad picture of himself Even the *K* scale was empirically derived Items in the *K* scale were selected on the basis of their ability to discriminate between clinically normal persons producing abnormal test scores and clinically abnormal individuals with abnormal scores, and between clinically abnormal persons with normal test scores and abnormal subjects whose test records were abnormal.

THE BEGINNINGS OF RESPONSE-SET THEORY

This line of investigation—the empirical identification of, and correction for, the opposite effects of "faking good" and "faking

bad"—constitutes one major approach to the stubborn resistance of personality tests, and the subjects who take them, to produce consistently predictive responses. Approaching the problem from a somewhat different perspective, Cronbach (1946, 1950) proposed that response bias stems from idiosyncratic test-taking habits of the subject—personal sets to respond to content-irrelevant aspects of the test. Among the various response sets he described are the tendency to assign uncommon meanings to such response categories as "agree," "uncertain," or "like," the tendency to guess, and acquiescence. Outstanding among these response sets is the tendency to acquiesce—to mark *true*, indicate agreement, or otherwise respond uncritically to test items by personal endorsement.

In Cronbach's conception of response sets can be discerned the beginnings of a theory of test-taking behavior seeking to take into account the meaning to the subject of the test, of the test items, and of being tested. Despite his perceptive recognition that response sets may reflect significant habits and attitudes of the subject, Cronbach unfortunately went on to place major emphasis on the removal of these sources of variance from tests. Thus, he fell back to the old position of the primacy of item content.

Acquiescence has been established as a major response determinant in the measurement of such personality variables as authoritarianism (Bass, 1955, Chapman & Campbell, 1957, Jackson & Messick, 1958). The basic method has been to show, first of all, that a given questionnaire—say the California *F* scale (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950)—has a large proportion of items keyed *agree* (or *true* or *yes*). Second, half the items are reversed, now being scored for disagreement. Correlations are then computed between the original and reversed items. Failure to find high negative correlations is, then, an indication of the operation of response acquiescence. In one study on the *F* scale, in fact, significant *positive* correlations—strongly indicative of an acquiescent tendency—were found (Jackson, Messick, & Solley, 1957). Another procedure is to employ the percentage of items keyed *true* on a given scale, correlating this index with measures known to reflect acquies-

cence—again, say, the *F* scale High positive correlations presumably indicate an agreeing tendency Still a third method is to use a large number of items varying widely in content, i e , in their psychopathological implications Then, employing a Likert scale to afford the subject several response alternatives, an overall mean can be computed for each subject which reflects his tendency to agree with a large number of items whose content is so heterogeneous that consistent endorsement on content grounds would constitute a blaze of contradictions This was the model followed by Couch and Keniston (1960) in their detailed analysis of the test behavior and personality of "yeasayers" and "naysayers "

In a thoughtful paper, Jackson and Messick (1958) drew a distinction between content and style in the measurement of personality Developing the argument in part from Allport and Vernon's (1932) well-known analysis of expressive behavior, they proposed that

> In personality theory a ubiquitous and fundamental distinction may be drawn between the interpretation of behavior in terms of (*a*) the *content* of "needs" and of cognitive stiuctures generally and in terms of (*b*) characteristic *styles* of response and action   One may legitimately ask not only *what* a person says or does (the particular content of his statements and actions) but *how* he acts (his characteristic *mode* oi *style* of expression) (p 243)

Applying this differentiation to the assessment of personality characteristics or attitudes, Jackson and Messick contended that both stylistic properties of *test items* and habitual expressive or response styles of *individuals* may outweigh the importance of item content The way an item is worded—its style of expression—may tend to increase its frequency of endorsement The tendency of an individual to acquiesce represents a personal expressive mode or style on personality inventories which may be generalized to other situations For example, recall the song in the musical "Oklahoma," "I'm just a girl who can't say 'no ' " Jackson and Messick concluded that in heterogeneous personality inventories like the MMPI consistent responses to content are extremely difficult to identify Specifically, one of the two major factors repeatedly identified in factor analyses of the

MMPI (Welsh, 1956, Wheeler, Little, & Lehner, 1951) may be best explained by an acquiescent response set or style (Messick & Jackson, 1961)

Whether one is disposed to accept an acquiescence interpretation of personality-test scores or whether some variant of the dissimulation or "faking good" hypothesis is more appealing, contemporary theorizing and research in personality assessment lean sharply toward the stylistic or response-set view Indeed, Berg titled a chapter in *Objective Approaches to Personality Assessment* (Bass & Berg, 1959) "The unimportance of item content" Berg's response-set proposal is a somewhat different one than those we have considered so far In Berg's view, the emphasis is on deviant rather than stereotyped or conventional responses Deviant behavior, he claims, is generalized or transsituational Thus, atypical personality-test responses—responses which deviate from the norm, which are statistically infrequent—can be expected to be associated with deviant or atypical behavior in other areas The Deviation Hypothesis, however, must be recognized as a starkly atheoretical approach to response styles, an approach which, as Norman (1963) has pungently argued, is of technical significance at best

Despite the multiplicity of approaches to the problem of response distortion, our brief review warrants the conclusion that an earlier naiveté—first an uncritical acceptance of the literal significance of item content and later a cynical view of the subject's honesty—has given way to somewhat more sophisticated and methodologically sound analyses In particular, the distinction between content and style and the emphasis on expressive aspects of the subject's test behavior suggest the meaningfulness of sets to respond in a particular fashion on personality questionnaires As we noted, Couch and Keniston (1960) have in fact shown that "yeasaying"—the undiscriminating tendency to agree—is widely general over different items and tests and stems from "a central personality syndrome" (p 173)

To anticipate our discussion of the testing situation in the next chapter, response sets may further be considered in an instrumental sense—as meaningful problem-solving behavior indicative of an individual's approach to being evaluated and re-

lated to his behavior in other situations In an able review, McGee (1962) has brought together and critically evaluated some recent attempts to demonstrate the personality characteristics of the various response styles Thus, a measure of theoretical concern is beginning to appear on the scene, supplanting the earlier reliance on the strictly empirical identification of response sets In this more recent work, relating response styles on personality inventories to behavior in nontest situations, the simple technical mastery of prediction is surpassed, and the possibility of understanding the meaning and generality of test behavior is greatly enhanced

In the next chapter we turn to the problem of a social-desirability response set, which vies with acquiescence as the major stylistic determinant on personality inventories The disposition to respond in a socially desirable manner is the conceptual point of departure of this book, and the quest for its motivational determinants is the essential *raison d'être* of the experiments to be described in the following chapters

2

Social desirability and the test-taking situation[1]

To prospective victims of the "black art" of personality testing W H Whyte (1956) gave the following advice

> When an individual is commanded by an organization to reveal his innermost feelings, he has a duty to himself to give answers that serve his self-interest rather than that of The Organization In a word, he should cheat To put it so baldly may shock some people—I was scolded severely by several undergraduate groups for giving just such advice But why be hypocritical? Most people instinctively cheat anyway on such tests Why, then, do it ineptly? When in doubt about the most beneficial answer to any question, repeat to yourself I loved my father and my mother, but my father a little bit more I like things pretty much the way they are I never worry much about anything (pp 179, 196–197)

Whyte's counsel, and his more elaborate crib for taking personality tests, were intended to afford a measure of self-protection from the entrenched values, norms, and predilections of The Organization, to which failure to conform might well result in failure to be hired or promoted

The Organization Man's attack on the bureaucratized use of tests to insure the selection of congenial servants of The Organization sought to brand such efforts as iniquitous and scientifically specious In another sense, however, Whyte's critique is a trench-

[1] Portions of this chapter are taken from Crowne and Marlowe (1960)

ant and suggestive social commentary on the pervasive attitudes of *individuals* toward tests, testing, and personal disclosure—attitudes which have plagued personality assessment from its very beginnings. Whyte highlights the resistance of many persons to revealing their beliefs and feelings and their penchant for favorable public self-evaluations.

The test-taking strategy which he commends is, then, one which a large number of people seem naturally to evolve for themselves. Its outstanding characteristic is the evaluation of oneself in socially desirable terms. Thus, on any item the choice of acknowledging or denying some trait, feeling, or symptom is made by deciding which alternative would result in the most socially favorable self-description.

But where does the problem really lie? In the subject's motives and attitudes toward being tested? This seems to be a likely possibility, and it is one to which we shall return later in this chapter. There is another approach, however: to take the characteristics of personality-test items as a point of departure and to ask, "What is it about test items that would cause subjects to endorse or to reject them on grounds other than content?"

THE SOCIAL DESIRABILITY OF PERSONALITY-TEST ITEMS

Several years ago, Edwards (1953a) took this approach and proceeded to demonstrate an intriguing relationship. He correlated judges' ratings of the social desirability of personality-test items with the probability of their endorsement. The correlation turned out to be .87: the more favorable the social-desirability rating of an item, the greater the likelihood of its endorsement under standard test-taking instructions. In his subsequent monograph, Edwards (1957) formally defined a new factor in the distortion of test responses as a dimension of social desirability. We have seen that earlier analyses of response distortion—Meehl and Hathaway's (1946) *K* scale and Cronbach's (1946, 1950) notion of response sets—focused on the *test behavior* of the individual. Social desirability, however, was intended to describe the tendency of *test items*

to elicit responses in the favorable direction Edwards (1957) contended that

> regardless of the multidimensional nature of personality statements, with respect to content, it is still possible to describe each one in terms of its position on the social desirability continuum The social desirability continuum appears to me to be the most important single dimension on which to locate personality statements The importance of the dimension rests in the fact that, if we know the position of a statement on it, we can then predict, with a high degree of accuracy, the proportion of individuals who will say, in self-description, that the statement does describe them (p 3)

Since the original work of Edwards, extensive efforts have been devoted to investigating the relationship between the social desirability of personality-test items and the probability of their being acknowledged by subjects in self-description Two models have been applied in the study of this relationship The first of these is Edwards' (1953a) basic procedure of correlating scaled social-desirability ratings with actual frequency of endorsement Cowen and Tongas (1959), Rosen (1956), and Wiggins and Rumrill (1959), among others, have applied this model to a variety of personality tests, reporting findings very similar to Edwards' original demonstration It has further been established that males and females rate the social desirability of personality-test items in a virtually identical fashion (Edwards, 1957) Moreover, Nisei (Fujita, 1956), hospitalized psychiatric patients and "Skid Row" alcoholic tuberculosis patients (Edwards, 1957), high school students from different socio-economic classes (Klett, 1956), and Norwegians (Edwards, 1957) produce social-desirability ratings dramatically similar to those of the undergraduate college students in Edwards' original study Thus, not only is it likely that items in most personality tests will be endorsed if they are independently rated as desirable and rejected if they are rated as undesirable, but there appears to be widespread agreement on what *is* socially desirable among groups presumably varying greatly in cultural background, social position, and personality

The second approach to investigating the effects of social desirability has been the development of rational or empirical

scales composed of items showing marked social-desirability properties These instruments then constitute independently derived predictors of biased test responses In addition to the social-desirability scaling of test *items*, Edwards (1957) also constructed such an index of the relative tendency of different subjects to produce socially desirable test responses

THE EDWARDS SOCIAL-DESIRABILITY SCALE

In the development of a social-desirability (*SD*) scale, Edwards drew a sample of 150 items from various MMPI scales (*F*, "plus-getting" and test validity, *L*, the Lie scale, *K*, the dissimulation or "faking good" and "faking bad" scale, and the Manifest Anxiety scale [Taylor, 1953]) and submitted them to ten judges instructed to give the socially desirable response The judges agreed unanimously on 79 items, which constituted the initial measure A subsequent version, incorporating the 39 items which best discriminated between high and low total scores, has since become the primary instrument in the measurement of social desirability by this procedure

When the Edwards *SD* scale was correlated with the validity, clinical, and derived scales of the MMPI and with a variety of other personality inventories, the expected pattern of results emerged Very high negative correlations were found between the *SD* scale and most of the clinical and derived scales, which are conventionally scored in a socially undesirable direction. Positive correlations were obtained between the *SD* scale and the *K* and *L* scales, both of which are keyed in the socially desirable direction The correlation with the *F* scale, in which the items are primarily keyed in the socially undesirable direction, was negative The pattern of these correlations can be seen in Table 2 (p 26) in the column headed Edwards *SD* scale These findings are taken from Crowne and Marlowe (1960), but they closely approximate the original results of Merrill and Heathers (1956) reviewed in Edwards' monograph

On those scales containing one group of items with obvious pathological implications, scored in the obvious direction, and another group of "subtle" items, the *SD* scale correlated nega-

tively and very highly with the obvious items and positively and less substantially with the subtle items The contrast between subtle and obvious items and their implications for socially desirable responding are illustrated by the following examples Both are from the Hysteria scale of the MMPI The first (obvious) is scored if endorsed, the second (subtle) if denied "I am troubled by attacks of nausea and vomiting" "I enjoy detective or mystery stories" Clearly, the obvious items are scored in the socially undesirable direction, while the subtle items are keyed for socially desirable responses or are ambiguous as to the culturally acceptable response

From the magnitude of these correlations it is apparent that the systematic variance remaining after the correlations with the *SD* scale have been accounted for is small indeed Evidently, then, there is little left in personality testing but social desirability Since the *SD* scale relates so highly to the MMPI and to a wide array of personality tests, the logical extension of the social-desirability hypothesis would be that the *SD* scale is a kind of abbreviated universal personality test, from which an individual's scores on most other personality instruments could be predicted with impressive accuracy In a paper entitled "A short form of the MMPI: The *SD* scale," Edwards and Walker (1961b) have advanced just this contention

> It is obvious that a subject's score on the *SD* scale provides a relatively sound basis for predicting his scores on the other 58 [clinical and derived] MMPI scales It thus appears possible to use the prediction equation to establish expectancy tables for the various MMPI scales With such tables available, individuals need only be given the 39-item *SD* scale, and their predicted scores on the other MMPI scales could be obtained from the expectancy tables On the basis of the results we have reported, it seems that we could be fairly confident that the scores obtained from the expectancy tables would correspond very closely to the scores that would have been obtained if the subjects had been given the complete MMPI Obviously, use of the expectancy tables would save considerable time for both the examinee and the examiner (p 486)

Subsequent research with Edwards' *SD* scale and with other similar instruments has been chiefly concerned with expanding the descriptive analysis of social desirability Factor-analytic methods have been employed to identify components of response-

set variance on such tests as the MMPI and to attempt to isolate acquiescence, social desirability, and deviation response sets.

Little attention has been given to the generality of the disposition to respond in a socially desirable manner in situations in which self- or item-evaluation is not the primary dependent variable. Two studies, however, are concerned with individual differences in the tendency to respond in a socially desirable fashion and with the prediction of behavioral differences between subjects scoring high and low on the *SD* scale. Edwards and Diers (1962) compared high- and low-scoring groups on the *SD* scale in the number of question mark or avoidance responses to pairs of items on the Edwards Personal Preference Schedule (1953b). Some of the pairs had high and others low social-desirability scale values. They reasoned that when a subject is required to choose one of a pair of statements both of which have high social-desirability scale values, the problem is analogous to an approach-approach conflict. On the other hand, having to choose between two socially undesirable statements represents an avoidance-avoidance conflict. In this study an opportunity was provided to avoid making a decision by a question mark or "cannot say" alternative. It was predicted that the avoidance-avoidance conflict would promote more "leaving the field" behavior, as represented by the question mark option. Edwards and Diers found a slight tendency for subjects with high scores on the *SD* scale to make more question mark responses than low-*SD* subjects to the avoidance-avoidance choices relative to approach-approach choices. Overall, high-*SD* subjects made significantly more no-choice responses, that is, they tended to avoid the problem of making choices between pairs of items closely matched in social desirability.

In an attempt to predict the effects of the social-desirability response set in a situation other than personality testing, Allison and Hunt (1959) investigated the relationship of the Edwards *SD* scale to aggressive responses to varying conditions of frustration as measured by a paper-and-pencil test. They defined three categories of frustrating situations: those in which aggression would be socially justified, those in which it would not, and those in an unspecified or ambiguous category. High-

SD subjects made significantly fewer aggressive responses in the unspecified condition than did low-*SD* subjects. Allison and Hunt found that subjects characterized by a tendency to evaluate themselves favorably on the *SD* scale gave fewer aggressive responses only under conditions in which the cultural definition of an acceptable response was not evident. They interpreted their findings to mean that these individuals are capable of expressing aggression under conditions in which the culture provides some indication of the appropriate response. This study differs from previous work with the *SD* scale in that subjects were not required to acknowledge or to deny maladjustive symptoms. The measure of aggressive response to frustration, however, was still self-evaluative rather than behavioral. As a consequence, it is difficult to draw meaningful conclusions about the motivational or dispositional properties of the *SD* scale in nontest situations.

There are major problems in interpreting each of the conceptualizations of test-irrelevant response determinants—"faking good" or "faking bad," acquiescence, the tendency to give deviant responses, and social desirability. One problem concerns whether the dispositions to agree, to dissimulate, or to respond in a socially desirable fashion are restricted to situations involving self-evaluation, or whether these propensities are reflections of more general personality characteristics of the individual. In other words, are the effects of the various response sets to be observed only in personality testing, or are there behavioral (nontest) correlates as well?

A second problem is how to interpret the correlations between response-set measures and personality inventories. For social desirability, correlations between the Edwards *SD* scale and such MMPI scales as Psychasthenia, Schizophrenia, and Manifest Anxiety reach values in the neighborhood of −.80. When we consider that there are substantial errors of measurement associated with these scales, it is evident that virtually all of the non-error variance can be accounted for by correlations of this magnitude. Such sizeable correlations lead to an awkward complexity in interpretation: either the tendency to respond in a socially desirable fashion is all that is being

measured, or social desirability and manifest anxiety, for example, amount to the same thing Neither is a palatable conclusion, and in neither case do we know any more about manifest anxiety or about a social-desirability response set These observed correlations are still shrouded in mystery, despite the fact that they are mediated in part by item overlap and by similarity in item content, the problem of interpretation remains

Although the amount of work has been voluminous and the conceptualizations have been profuse, it is not yet understood why people acquiesce, dissimulate, or respond in a socially desirable manner Nor if one were to take the view that the distortion of test responses represents motivated behavior is it clear in what kinds of nontest situations the motives affecting test-taking behavior would be influential Still another problem is the lack of understanding of the relationships among the various measures and interpretations of acquiescence, dissimulation, deviation, and social desirability The recent literature on test-taking attitudes has witnessed a sharp debate between the proponents of an acquiescence factor as the primary determinant of test-response distortion (Couch & Keniston, 1961) and advocates of a unidimensional social-desirability hypothesis (Edwards & Walker, 1961a, Edwards & Walker, 1961c, Taylor, 1961) Two additional possibilities have been suggested by Jackson and Messick (1961) and Wiggins (1962) On the basis of a factor analysis Jackson and Messick argue that social desirability and acquiescence can be regarded as orthogonal, and they further suggest that acquiescence is differentially evoked by items with high, middle, and low social-desirability scale values Wiggins, on the other hand, contends that the most advantageous conceptual point of departure is represented by deviant-true and deviant-false response sets which derive from Berg's (1955) Deviation Hypothesis

The prevalent conception of social desirability reflects a virtually exclusive concern with response distortion in psychometric situations, little research has gone beyond investigations of the social-desirability scalability of test items or the correlation of personality tests with independently derived social-desirability scales Indeed, the restriction of interest to self- or item-

evaluation in the analysis of test-taking attitudes and dispositions is characteristic of research in this area, with the possible exception of acquiescence (Couch & Keniston, 1960, Jackson & Messick, 1958) McGee (1962) pointedly comments on the paucity of studies relating response styles to observable behavior

The primary research strategy on the response-set problem, as exemplified by the work on social desirability, appears to have been guided in its initial stages by the implicit assumption that sets to respond in a particular fashion constitute sources of error variance Recently, however, there has been a growing recognition of the probability that response styles may reflect more general and important behavioral characteristics It is nevertheless true that response styles are not anchored in personality theory, they continue to be regarded as primarily testing phenomena

To cap our argument up to this point, response-bias research has been characterized by two major weaknesses First, the inherent confounding of test-item content with determining stylistic tendencies has not clearly been recognized The descriptive analysis of responses to personality tests, however sophisticated and refined, offers little likelihood of revealing the motives or personality characteristics of individuals who display a set to respond in a particular fashion It is difficult if not impossible to draw viable conclusions about the personality characteristics of people who display a social-desirability response set from their scores on various personality scales, since these measures are themselves dramatically saturated by the very tendency to give socially desirable responses Thus, if an individual's test responses are characterized by a consistent tendency to endorse the socially desirable alternatives whatever the content of the items, it is not possible to conclude anything about his personality characteristics from his test scores They simply reflect his test-taking strategy Or, to put it another way, personality-test responses cannot be used at the same time to establish the existence of a social-desirability response set and as a personality description Although statistical methods such as factor analysis bring a high level of sophistication to the problem,

they do little to resolve this fundamental circularity What is required is the attempt to relate response-set measures to theoretically relevant and methodologically independent *behavioral* criteria The second weakness lies in the failure to recognize the testing situation as a person-situation interaction in which the individual's goals and expectations are variables of major importance

These considerations suggest the need for a reformulation, and it is to a recasting of the particular problem of the social-desirability response set that the research to be described in the following section is addressed

DEVELOPMENT OF A NEW SOCIAL-DESIRABILITY SCALE

In the construction of measures of individual differences in social-desirability response bias, we have seen that the usual procedure has been to draw a variety of items from a personality inventory and to have judges rate them as to their social desirability or undesirability Those items which the judges unanimously regard as being highly socially desirable (or undesirable) statements to attribute to oneself constitute the social-desirability scale Persons who endorse socially desirable items and reject socially undesirable ones are said to be demonstrating a social-desirability response set

It must be remembered, however, that the items were drawn from clinical scales and may be characterized by their pathological *content* as well as by social desirability Typical items in such a scale are "Criticism or scolding hurts me terribly," and "I feel anxiety about something or someone almost all the time" Rejection of these items may imply a social-desirability response set, but it is also conceivable that subjects responding *false* may simply not feel generally anxious or be traumatized by criticism, there is no way of telling the difference The probability of occurrence of symptoms represented in MMPI items (and incorporated in the Edwards and most other *SD* scales) in a college undergraduate population is undoubtedly low The achievement of high social-desirability scores, then,

may reflect only the low frequency of maladjustive symptoms in this population and not the disposition of subjects to present themselves in a favorable light

To understand the meaning of scales constructed according to this model, it is crucial to recognize that such instruments do not indicate the motivated willingness or unwillingness of subjects to acknowledge maladjustive symptoms but simply whether subjects *do* admit to symptoms or not If one assumes that the goals of subjects in psychometric situations have an important influence on test responses, it is essential to be able to discriminate between the effects of item content and the influence of motives

To conceptualize the matter somewhat differently, imagine a class of statements that one might make about oneself which are defined by two major attributes First, they are "good," culturally sanctioned things to say about oneself, and second, they are probably untrue of most people Better yet would be a balanced scale composed half of culturally acceptable but probably untrue statements and half of true but undesirable statements This is the model for the development of the Marlowe-Crowne Social-Desirability (*M-C SD*) scale The antecedent of this rationale will be recognized in the prior work of Meehl and Hathaway (1946) on the Lie scale and in the even earlier *Studies in Deceit* of Hartshorne and May (1928) Remarkably, Hartshorne and May developed a scale for children to measure lying to win approval, an instrument which is very similar to the Lie and *M-C SD* scales Children achieving high scores on the "social approval" questionnaire were regarded as "pious frauds", and in the same sense persons getting high scores on the MMPI Lie scale were considered to be deliberately presenting a whitewashed self-appraisal

In distinction to Meehl and Hathaway and Hartshorne and May, however, the interpretation of responses on the *M-C SD* scale is somewhat different It is not necessary to assume either that subjects who acknowledge the "good" items and reject the "bad" items on the *M-C* scale are accurately describing how they actually behave or that they are consciously lying and that their responses represent witting and deliberate deceit The

meaning of high scores will be considered later in much greater detail, at this point it is sufficient to say that individuals who depict themselves in very favorable terms on the scale can be understood as displaying a social-desirability response set

Scale Construction

A number of current personality inventories were consulted in order to devise a set of items for the new scale To be included, an item had to meet the criterion of cultural approval and yet be untrue of virtually all people, and have minimal pathological or abnormal implications A set of 50 items meeting these criteria was submitted to ten judges for social-desirability ratings[2] The judges were instructed to score each item in the socially desirable direction from the point of view of college students, using true and false response categories Unanimous agreement was obtained on 36 items and 90 per cent on 11 more These 47 items constituted the preliminary form of the scale

A major aim in the development of the *M-C* scale was the elimination of items with psychopathological content As a test of our success in achieving this objective, both the *M-C* and the Edwards *SD* scales were submitted to ten additional judges (who were as experienced as the judges above) for ratings of the degree of maladjustment implied by socially undesirable responses to the items A 5-point scale, ranging from extremely well adjusted (1) to extremely maladjusted (5), was employed for this purpose The mean rating for all the items in the *M-C* scale was 2 8, slightly below the midpoint of the rating scale (implies neither good nor poor adjustment) The mean rating for the Edwards items was 3 9, indicating that the judges considered socially undesirable responses on this scale to be definitely indicative of maladjustment A t test of the significance of the difference between these means was 15 27, significant well beyond the 0001 level.

---

[2] Serving as judges were both faculty members and graduate students in the Department of Psychology at The Ohio State University.

The preliminary version of the scale was then administered to 76 students in two introductory psychology sections, and an item analysis was undertaken There were 33 items discriminating at the 05 level or better between high and low total scores Of the 33 items, 18 are keyed true and 15 false, making an acquiescence interpretation highly improbable These 33 items constitute the final form of the *M-C SD* scale The scale is shown in Table 1

TABLE 1 The Marlowe-Crowne Social-Desirability Scale

Personal Reaction Inventory

Listed below are a number of statements concerning personal attitudes and traits Read each item and decide whether the statement is true or false as it pertains to you personally

1 Before voting I thoroughly investigate the qualifications of all the candidates (*T*)
2 I never hesitate to go out of my way to help someone in trouble (*T*)
3 It is sometimes hard for me to go on with my work if I am not encouraged (*F*)
4 I have never intensely disliked anyone (*T*)
5 On occasion I have had doubts about my ability to succeed in life (*F*)
6 I sometimes feel resentful when I don't get my way. (*F*)
7 I am always careful about my manner of dress (*T*)
8 My table manners at home are as good as when I eat out in a restaurant (*T*)
9 If I could get into a movie without paying and be sure I was not seen, I would probably do it (*F*)
10 On a few occasions, I have given up doing something because I thought too little of my ability (*F*)
11 I like to gossip at times (*F*)
12 There have been times when I felt like rebelling against people in authority even though I knew they were right (*F*)
13. No matter who I'm talking to, I'm always a good listener (*T*)
14 I can remember "playing sick" to get out of something (*F*)

TABLE 1 THE MARLOWE-CROWNE SOCIAL-DESIRABILITY SCALE (*Continued*)

15 There have been occasions when I took advantage of someone (*F*)
16 I'm always willing to admit it when I make a mistake (*T*)
17 I always try to practice what I preach (*T*)
18 I don't find it particularly difficult to get along with loud mouthed, obnoxious people (*T*)
19 I sometimes try to get even, rather than forgive and forget (*F*)
20 When I don't know something I don't at all mind admitting it (*T*)
21 I am always courteous, even to people who are disagreeable (*T*)
22 At times I have really insisted on having things my own way (*F*)
23 There have been occasions when I felt like smashing things. (*F*)
24 I would never think of letting someone else be punished for my wrongdoings (*T*)
25 I never resent being asked to return a favor (*T*)
26 I have never been irked when people expressed ideas very different from my own (*T*)
27 I never make a long trip without checking the safety of my car (*T*)
28 There have been times when I was quite jealous of the good fortune of others (*F*)
29 I have almost never felt the urge to tell someone off (*T*)
30 I am sometimes irritated by people who ask favors of me (*F*)
31 I have never felt that I was punished without cause (*T*)
32 I sometimes think when people have a misfortune they only got what they deserved (*F*)
33 I have never deliberately said something that hurt someone's feelings (*T*)

Reliability

To determine the reliability of the scale, both internal consistency and test-retest coefficients were obtained Using the Kuder-Richardson formula 20, the internal consistency coefficient for the final form of the scale was 88 Fifty-seven sub-

jects[3] took the scale on two occasions separated by a one-month interval A test-retest correlation of 88 was obtained These correlations indicate that reliability was very satisfactorily achieved

Correlations with Other Scales

To investigate the response-set properties of the new questionnaire and to compare its performance with the Edwards scale, 39 students in an abnormal psychology class were given the *M-C* scale and the MMPI Pearson product-moment correlations were computed between the two *SD* scales[4] and the following MMPI and derived scales *K*, test-taking attitudes, *L*, lie, *F*, validity and test-taking attitudes, *Hs*, hypochondriasis, *D*, depression, *Hy*, hysteria, *Pd*, psychopathic deviate, *Pa*, paranoia, *Pt*, psychasthenia, *Sc*, schizophrenia, *Ma*, manic, *Pr*, prejudice (Gough, 1951), *St*, status (Gough, 1948), *Es*, ego strength (Barron, 1953a), *MAS*, manifest anxiety (Taylor, 1953), *A*, anxiety (Welsh, 1956), *R*, repression (Welsh, 1956)

Table 2 presents the correlations of the two *SD* scales with these 17 MMPI validity, clinical, and derived scales It is at once apparent that there are uniformly higher correlations between the Edwards *SD* scale and the various MMPI scales than between the *M-C* scale and these MMPI variables A general trend, consistent with previous research, is found in the positive correlations of the two *SD* scales with the *K* and *L* validity scales of the MMPI and the negative correlation with *F*, interpreted as an index of "plus-getting" Negative correlations appear with most of the clinical scales Four clinical scales correlate highest with both *SD* scales, with the single exception of *D*, which correlated − 27 with the *M-C SD* scale *Sc*, *Pd*, *Pt*, and *Hs* Two of these four, *Sc* and *Pt*, are considered to be among the most "pathological" of the clinical scales Dividing

[3] Thirty-one of these subjects were from the original sample, the remainder, from later experimental samples, repeated the scale under comparable conditions

[4] The Edwards scale was scored from the MMPI

TABLE 2 CORRELATIONS BETWEEN THE SOCIAL-DESIRABILITY SCALES AND VARIOUS MMPI SCALES FOR 37 MALES AND FEMALES

	M-C SD Scale			Edwards *SD* Scale		
MMPI Scales	Total Scale	Subtle Items	Obvious Items	Total Scale	Subtle Items[a]	Obvious Items[a]
K	.40 *			65 †		
L	54 †			22		
F	– 36 *			– 61 †		
Hs	– 30			– 62 †		
D	– 27	14	–.37 *	– 72 †	.33 *	– 78 †
Hy	.15	32	–.23	09	54 †	–.71 †
Pd	– 41 †	– 21	– 46 †	– 73 †	.27	– 85 †
Pa	21	.37 *	– 18	– 02	06	– 72 †
Pt	– 30			– 80 †		
Sc	– 40 *			– 77 †		
Ma	– 24	04	– 32	– 42 *	40 †	– 53 †
Pr[b]	– 27			– 58 †		
St[b]	.16			14		
Es	17			.46 †		
MAS[c]	– 25			– 75 †		
A[c]	– 23			– 61 †		
R[c]	.28			07		

* Significant at the 05 level
† Significant at the 01 level
[a] From Edwards (1957).
[b] $N = 36$
[c] $N = 34$

the items on five of the clinical scales (*D, Hy, Pd, Pa,* and *Ma*) into those scored in a subtle rather than an obvious direction (Weiner, 1948), the correlations of the obvious items with the *M-C SD* scale are consistently negative and more substantial than the correlations with the total scales Relationships with the subtle items tend to be low and positive These findings parallel those with the Edwards scale that we considered earlier

In contrast to the Edwards scale, which correlates with certain of the MMPI clinical scales (e g, *Pt*, *Sc*, and *MAS*) in a degree approaching the asymptotic value of the reliabilities of the separate tests, the correlations of the *M-C* scale with the MMPI are much lower We found the predicted tendency for those subjects achieving high scores on the new measure to achieve low (nonpathological) scores on a number of the MMPI clinical and derived scales, but the correlations obtained were low to moderate, indicating that the *M-C SD* scale is accounting for some, but not all, of the variance These smaller correlations would be predicted if one were to view social desirability as accounting for a part of test-response variance, but not all or most of it The magnitude of these correlations more accurately reflects the amount of response-set variance which might plausibly be attributed to differences in the *need* to give socially desirable responses

What kinds of motives, then, would impel subjects to endorse culturally acceptable statements on the *M-C* scale and to tend to achieve more or less "adjusted" scores on such personality tests as the MMPI? Our initial interpretation of socially desirable responding on the new scale was that people conform to social stereotypes of what is good to acknowledge concerning oneself in order to achieve approval from others The *M-C SD* scale was, thus, an indirect measure of a need for approval At this provisional stage, the construct of need for approval implied that (*a*) people differ in the strength of their need to be thought well of by others, and (*b*) for those whose need is higher, we could assume a generalized expectancy that approval satisfactions are attained by engaging in behaviors which are culturally sanctioned and approved (and by avoiding those responses which are not) This formulation appeared to account fairly well, theoretically, for test-taking behavior That is, it is simply not considered desirable in the contemporary social milieu to indicate on a test that one is anxious, frustrated, unhappy, and beset by all sorts of strange thoughts and impulses It is not consistent behavior if one is dependent upon the acceptance, recognition, and approval of others

THE TESTING SITUATION

Implicit in this formulation is the view that the testing situation is similar to other socially evaluative situations and that concepts useful in the explanation of social behavior in other evaluative contexts can be profitably employed to understand the motives and social forces affecting test responses

In the history of personality assessment, little concern has been given to the testing situation as a meaningful social context which might itself influence an individual's behavior In the effort to validate tests for both applied and theoretical purposes, the regnant goal has been to predict an individual's status on some criterion measure rather than to understand his behavior in a given social setting (the test situation) and to attempt to relate this behavior and its determinants to behavior in other situations Mainly, then, the preponderance of effort in personality testing has been directed toward the determination of class membership, by means of what Wiggins (1962) has referred to as criterion group analysis As an example, the items on the MMPI were selected for their power to discriminate between individuals belonging to clinical criterion groups (e g, hysterics, schizophrenics, or paranoids) and nonhospitalized persons not belonging to such groups

There is an extensive literature on the effects of role playing and faking, instructional sets, and variations in the characteristics of the examiner and the testing context in the distortion of responses to both structured and projective personality tests, but it still seems to have been assumed that class membership is a more powerful determinant of test responses than motives to distort or the social structure of the test situation Derived from this is the further assumption that any response determinants which obscure an individual's class membership constitute error variance which can potentially be identified and controlled or eliminated The forced-choice method in test construction was introduced precisely to minimize these obscuring influences. The Personal Preference Schedule is a familiar illustration of this method in its attempt to control for the influence

of social desirability by requiring the respondent to choose between pairs of items matched for social desirability

By implication it follows that the nature and meaning of the test situation are of limited significance, contributing only to testing error Thus, research on faking (Cofer, Chance, & Judson, 1949), investigations of social-desirability-instructed groups (Wiggins, 1959), and crude analyses of the meaning of the test situation for such clinical groups as paranoids (as exemplified by subtle or "trapping" scales) and in particular contexts such as industrial screening are only of importance in terms of efforts to make tests less susceptible to distortion

In his critique of personality testing, Whyte (1956) remarked that "The dehumanized literature of the field gives little hint of the highly personal overtones that mark the meeting of interviewer and interviewee" (p 186) In the confrontation of examiner or experimenter and subject both the motives and anticipations of the subject and the characteristics of the situation itself contribute to a highly charged encounter in which the subject's responses are most unlikely to escape unaffected Whyte, of course, was concerned about the inappropriate invasion of personal privacy in the screening of candidates to meet The Organization's purposes Equally compelling, however, is the application of personality tests in clinical situations where, for example, an individual's admission to, retention in, or discharge from a psychiatric hospital may depend, entirely or in part, upon his test responses Acceptance of a patient for psychotherapy and the utilization of tests to select those patients who are deemed likely to profit therefrom is analogous in the situational pressure exerted on the individual

A number of writers (Barthel & Crowne, 1962, Orne, 1959, Riecken, 1962, Rosenberg, 1961) have commented on the social forces operative in the psychological experiment or the testing situation If psychological experimentation is not regarded as outright sorcery, it is at least widely believed that psychologists are concerned with the detection of dark personal secrets of which the subject is himself unaware Certainly, an experimental subject is likely to be concerned about the consequences of his behavior, the very nature of the situation itself as it is

often defined for him by the culture tends to create such a concern How the experimental or testing situation is perceived by the subject may be augmented or more sharply defined by instructions, induced sets to respond in a particular manner, prior experimental manipulations (as, for example, in the stressful procedures used by Lindzey [1950a, 1950b]), or variations in the character of the situation itself

The sheer fact of being tested, then, is likely to arouse the subject's needs and anticipations related to social evaluation While the immediate setting in which a test is given can produce powerful effects on test responses, we are here concerned with the ways in which an individual's past experiences dispose him to interpret evaluative situations and with the characteristic behaviors which follow from his interpretations It is not enough merely to understand the testing situation as it is culturally defined or to control it operationally in experimentation; we must be able as well to conceptualize the meaning of testing for a given individual In other words, to know anything about an individual's personality characteristics from tests we have to know what it means to him to be tested This is no less a concern of personality theory than the content of the test items themselves, and indeed it appears to represent an even thornier and more complex issue

NEEDS, EXPECTANCIES, AND THE SITUATION

In a perceptive analysis Rotter (1960) has applied the concepts of Social Learning Theory to an understanding of the testing situation, the behavior to be predicted from tests, and the intricate relationships among tests, the testing situation, and behavior The major outlines of Social Learning Theory (Rotter, 1954) will be detailed in a later chapter devoted to an attempt to incorporate the construct of need for approval within the larger domain of a personality theory For present purposes we shall be concerned with the three broad classes of variables in this theory that determine behavior In Social Learning Theory behavior is a function of the person's goals or needs (the relative value of, or preference for, reinforcing

events), his expectancies that his goal or goals will be achieved or frustrated as a result of certain behaviors in a given situation, and the meaning of the situation for him (that is, the cues it provides or the expectancies it arouses regarding the probable outcomes of alternative behaviors)

Needs

What can be said of an individual's needs or goals in a testing situation? It is important to distinguish between (1) the prediction of some subsequent behavior from needs as assessed by tests and (2) the effect of an individual's goals on his test responses themselves The latter problem is the crucial one for an understanding of the testing situation

Motives or goals which could exert a powerful influence on test responses would include at least the following achievement goals—needs to attain success in an academic, social-recognition, or competitive business situation, needs to win approval and affection from others, and dependence—needs for help, protection, and succorance Crowne (1959) investigated social role preferences as they are related to self-accepting or self-rejecting test responses His measure of role preference was an instrument designed to assess the relative strengths of the need to be perceived as free from any inadequacies and the need to be regarded as self-critical or self-deprecatory It was found that subjects at the extremes of either role preference tended to deviate from the group norm in self-evaluative behavior Further, preference for the self-devaluative role was significantly related to needs for protection-dependency and love and affection The denial-of-inadequacy role preference was significantly associated with dominance and independence needs

The work on need Achievement and other motives by McClelland, Atkinson, Clark, and Lowell (1953) and Atkinson (1958), in which need strength is assessed from fantasy material, illustrates the influence of motives on test responses in relatively ambiguous or unstructured situations It is well to recognize, however, that an individual's goals regarding *this test situation itself* may affect his responses to projective stimuli, as indeed

McClelland et al have found If such motives regarding the testing situation itself are influential, they are inextricably confounded with the motive presumably being assessed by the fantasy productions Another problem in these studies is that it is not known in what way nAchievement might influence an individual's test responses in other self-evaluative situations

Other research has shown that it is possible to arouse motives experimentally and to demonstrate a relationship between these induced needs and subsequent test responses The well-known studies of Lindzey referred to earlier, in which severe frustration and the arousal of aggression resulted in changes in projective test responses, are cases in point Clark (1952) and Mussen and Scodel (1955) demonstrated the effects of prior sexual arousal, which was accomplished by the display of nude pictures to male subjects, on the number of sexual responses on the Thematic Apperception Test However, the possible influence of the subject's expectancy that sex and sexual responsiveness were being evaluated was not controlled in these studies

A person's needs may dispose him to present himself in a particular light in a testing situation If it is important for him to gain approval or acceptance, to deny inadequacies, to obtain dependency gratifications, or to achieve recognition or status, we may anticipate that his test responses will be bent to serve these aims As Rotter (1960) put it, "What we call faking is only our recognition of the fact that the subject is taking the test with a different purpose or goal than the one the experimenter wants him to have" (p 308)

Expectancies

Expectancies regarding the evaluative consequences of behavior constitute another major source of influence on test responses A considerable portion of the research on test dissimulation is, in fact, concerned with manipulating the subject's expectancies regarding appropriate or inappropriate behavior in the test situation—that is, his expectancies of what will happen if he behaves in a given way Although not explicitly formulated in these terms, much of the research experimentally manipulating

test responses by means of instructional sets, situational variations, and the induction of success, failure, or other task orientations, can be understood in terms of the effect on the subject's interpretation of the situation and his anticipations regarding the outcomes of alternative behaviors. In projective testing, the affective tone (Birney, 1958) and cue characteristics of the picture (Haber & Alpert, 1958) have been shown to be significantly related to achievement themata and imagery. The test stimuli themselves thus provide the subject with cues about the kind of behavior which is expected. Even more clearly, achievement imagery seems to be increased when the subject is instructed that the measure is an achievement or intelligence test (McClelland et al., 1953). Similarly, affiliation-related responses to the TAT were increased when this measure was preceded by a sociometric test in which fraternity brothers rated one another (Atkinson, Heyns, & Veroff, 1954). Rather than arousing motive strength, these procedures appear to have altered the subject's *expectancies* regarding outcomes of social acceptance or rejection. Crandall (1951) found that failure on physical skill tasks resulted in lowered expectancies of success assessed by a TAT type of measure.

The indirect relationship of fantasy measures and overt behavior further illustrates the importance of expectancies. Studies by Kagan (1956) and Mussen and Naylor (1954) demonstrated that expectancies of punishment for aggressive behavior significantly influenced the relationship between aggressive themes on the TAT and overt aggressive behavior. Specifically, those subjects having a high expectation of punishment for behaving in an aggressive fashion tended not to display overt aggressive behavior even when their TATs contained aggressive themes. Atkinson (1958) and Atkinson and Reitman (1956) have pointed out the need to assess subjective probabilities or expectancies in the prediction of achievement-related and risk-taking behavior. More recent studies by Atkinson, Bastian, Earl, and Litwin (1960) and Atkinson and Litwin (1960) reveal the enhanced prediction of probability preferences and risk-taking behavior when measures of both motive strength

and expectancy are obtained Although they do not directly bear on the influence of expectancies on test responses, these investigations strongly suggest that goals or motives constitute only one determinant of behavior and that expectancies regarding the outcomes of behavior must, to achieve a satisfactory level of prediction, be taken into account

The Situation

The third behavioral determinant in Social Learning Theory is the psychological situation itself The psychological situation, or the "life space" in Lewin's earlier conceptualization, refers to the meaning of a given behavioral context for an individual The nature of a situation for a given individual is defined by his expectancies for the outcomes of alternative behaviors in that situation As Rotter (1954, 1955) has pointed out, the significance of a particular setting can be determined by assessing the individual's history of reinforcement of alternative behaviors in similar past situations In more novel contexts, it is clear that the person will largely have to learn what to expect and how to contend with the forces confronting him by actually behaving in that situation Also, generalized and trans-situational expectancies, representing attitudes toward oneself and others acquired in the course of many different situations, may be applied, appropriately or inappropriately, in a new and different situation Self-confidence, a generalized anticipation of failure, and a belief that one is essentially a powerless victim of luck, fate, or impersonal social forces (Liverant & Scodel, 1960, Rotter, Seeman, & Liverant, 1962) are examples of such expectancies generalized across situations

To determine the meaning of a test situation structured interviews might be employed to assess the individual's expectancies in the immediate setting and his past experience in self- and socially evaluative situations In a study reported in a later chapter, we used a brief interview to determine subjects' expectancies, or "evaluation apprehensions," as Rosenberg (1961) has denoted them, concerning the meaning of an experimental

situation. In sum, the individual definition of psychological situations involves the assessment of the subject's expectancies regarding the goals to be attained and the probabilities of their attainment.

A second means of defining psychological situations, perhaps somewhat less precise, is to determine their meaning for a given social group or, where appropriate, for a larger culture. The social description constitutes the meaning of a particular situation. The analysis of a social group, institution, or culture involves in part describing the meaning of a variety of situations as they are normatively defined within the group. Most individuals belonging to a particular group will tend to hold beliefs and attitudes in common, and such "social definitions" of situations can be employed to predict the behavior of members of the group. Edwards (1953a, 1957) used a variant of this method when he determined the social meaning of personality-test items by requiring subjects to rate how socially desirable they would be as attributes of other persons. Whyte's critique is an example of social criticism applied to personality testing and can be taken as a reflection of the attitudes of many persons in our society toward tests, testing, and the social purposes, nefarious and otherwise, to which personality testing is put.

Even a cursory review of the literature of personality testing reveals the paucity of effort devoted to an understanding of the social and personal significance of formal self-evaluation. The immense difficulties encountered in the attempt to predict meaningful social behavior from tests bear eloquent and frustrating testimony to the intricacies of the interaction of the subject and the testing situation.

The motivational conceptualization presented in this chapter is focused on the goals and expectancies brought into play in self- and socially evaluative situations. The construct of need for approval was formulated in an effort to take account of one motivational determinant of test-taking behavior. The *M-C SD* scale, which we have proposed as a measure of the approval motive, reflects a self-evaluative style brought into play in the testing situation.

Individual differences in the need for approval, however, would imply behavioral differences in other situations, and we turn now, in Parts II and III, to a series of experiments designed to test the validity of the concept—a series of experiments predicting behavioral differences in a variety of nontest situations

part II Experimental studies in compliance, influencibility, and conformity

3

Implied demand and conformity[1]

People differ strikingly in the tendency to describe themselves in improbably favorable terms Given this fact about an individual—the socially desirable public evaluation of himself—one can predict the general tenor of his responses on any number of personality indices This, then, appears to be his self-evaluative idiom The last chapter developed the proposition that individuals differ in the strength of their motives to characterize themselves in socially desirable terms The set to respond in a socially desirable manner specifically reflects a need for social approval

If one is strongly motivated to seek the approval of important others, it follows that not only should his public evaluation of himself be bent to serve this aim but other behaviors should reflect it as well In a situation confronting the approval-dependent individual with a clearly implied demand for conventional, polite, acceptable behavior, our theory would predict compliance or conformity

To observe the different responses of persons high and low in the need for approval to perceived situational demands requires a situation in which the demand for socially acceptable behavior is unambiguous and in which, to comply with that demand, subjects' own views or attitudes must be suppressed

[1] This chapter is adapted from Marlowe and Crowne (1961)

Suppose subjects were given a very boring, repetitive, dull task to perform for as protracted a period as possible Hopefully, subjects would experience frustration if they were unable to leave the situation, as a result, they ought to develop negative attitudes toward the task and the experimenter responsible for their plight Were this boring task to be presented by an experimenter who played to the hilt his role as university professor, authority figure, and omniscient psychologist, the conflict between the subject's inescapably real attitudes and the implied demand of such an authoritative and important other should be an especially salient one Persons with a high need for approval should yield, molding their attitudes toward the task and the experimenter to conform with what he must want to hear To create this kind of situation a task previously used by Festinger and Carlsmith (1959) seemed ideally suited It simply requires the subject to pack spools of thread in a small box for a very long time

To define social desirability as behavior motivated by a need for approval and the expectancy that approval can be attained by behaving in culturally acceptable ways might seem to overlap with conformity Certainly, a relation between the need for approval and conformity would be expected, but the two concepts can be differentiated by regarding the need for approval as a motivational variable, whereas conformity is best conceptualized as a description of a certain class of behaviors Thus, predicting a relationship between socially desirable test-taking behavior and conformity assumes that conformity constitutes a behavioral category likely to be engaged in by persons seeking to gratify a need for approval While this definitional distinction seems reasonable enough, we can still ask whether the two concepts differ in their utility for predicting the same behavior As a test of this the Independence of Judgment scale (Barron, 1953b), a paper-and-pencil measure of conformity, was included in the experiment

Finally, since our conceptualization and its derived test differ from other definitions and measures of social desirability, the same results would not be expected from other social-desirability

scales Accordingly, Edwards' *SD* scale was incorporated in the design of the experiment to determine its ability to predict the favorability of attitudes toward spool packing

EXPERIMENTAL METHOD

Each subject appeared individually for his experimental appointment, was met by the experimenter, and conducted to the small room in which the experiment was run He was seated at a table directly opposite the experimenter, who introduced himself in a "professional" and rather aloof manner as follows

My name is Dr ——— I'm a psychologist and I'm conducting an experiment on measures of performance Before we get started on the experiment, I would like you to fill out these questionnaires Sign your name to all of them

The subject then completed the *M-C* scale Upon its completion he filled out the Barron Independence of Judgment scale The Barron scale is a questionnaire which has previously been shown to be valid in the experimental discrimination of conformers and nonconformers (Barron, 1953b, Tuddenham, 1959) A typical item is, "It is easy for me to take orders and do what I am told" Following this, approximately half the subjects completed the Edwards *SD* scale[2]

When the last questionnaire was finished, the experimenter, in a very businesslike manner, gave these instructions

Now for the experiment itself The materials are this box and the twelve spools I want you to take these spools, one at a time, and place them in the box When you are finished, empty the box and refill it, one spool at a time Continue to fill and empty the box until I tell you to stop Use one hand and work at your own preferred speed

[2] The remainder could not take the measure due to scheduling difficulties Actually, half our subjects packed spools first, answered the four questions, and then completed the inventories, the other half completed the inventories first and then were administered the task and the four questions The instructions were modified according to the order of presentation Subjects in these two conditions did not differ on means or variances on any of the measures, and the data were therefore analyzed without regard for order of presentation

The subject then packed and unpacked the spools for 25 minutes During this period the experimenter held a stopwatch, pencil, and pad, pretending to be busily engaged in timing and making notes on the subject's performance After 25 minutes the experimenter interrupted the subject and, abandoning his previous aloofness, assumed a breezily assured air

Okay, that's all we have in the experiment itself I hope you enjoyed it You get a chance to see how you reacted to the task and so forth I would like to know what your personal reactions are to the task and the experiment Would you answer this questionnaire

The questionnaire now handed to the subject contained the following four questions

1 Was the task interesting and enjoyable? Would you rate how you feel about the task on the scale below, where −5 means extremely dull and boring, +5 means the task was extremely interesting and enjoyable, and 0 means the task was neutral, neither interesting nor uninteresting
2 Did the experiment give you an opportunity to learn about your abilities and skills? Rate how you feel about this on a scale from 0 to 10, where 0 means you learned nothing, and 10 means you learned a great deal
3 From what you know about the experiment and the task involved in it, would you say the experiment was measuring anything important? That is, do you think the results may have scientific value? Rate your opinion on this matter on a scale from 0 to 10, where 0 means the results have no scientific value or importance, and 10 means they have a great deal of value and importance
4 Would you have any desire to participate in another similar experiment? Rate your desire to participate in a similar experiment, again on a scale from −5 to +5, where −5 means you would definitely dislike to participate again, +5 means you would definitely like to participate again, and 0 means you have no particular feeling about it one way or the other

Our subjects were 57 male undergraduate students enrolled in introductory psychology classes at The Ohio State University and at the University of Kentucky They participated in this and other experiments to fulfill a course requirement Typically, they sign up on a bulletin board where the sheets for the available experiments are posted Experiments are usually designated by a code number and an innocuous name to avoid biased

sampling This experiment was titled, "Measures of performance"

Immediately after the conclusion of the experiment subjects were informed of its purpose We might note here that in each of the experiments reported in the following chapters we apprised our subjects of the nature and the point of the deceptions practiced upon them Our concern was to avoid having subjects leave feeling more apprehensive or angry than they did when they came in

RESULTS

The major prediction was that individuals with a high need for approval would express more favorable attitudes toward the spool-packing task than subjects with a weaker approval need To test this, scores on the *M-C* scale were dichotomized at the overall mean (14 93),[3] giving a high-need-for-approval group of 30 subjects and a low-need-for-approval group of 27 subjects [4] The differences between the mean ratings given to the four questions by the high and low groups were then tested for significance by *t* tests Table 1 shows the results of this analysis

The differences in ratings on each of the four questions are all in the predicted direction, with the high-need-for-approval group consistently expressing more favorable attitudes than the low group Moreover, the evaluations of the high group are above the midpoint of the scale on each of the four questions By con-

[3] In this and in each of the experiments to follow high- and low-need-for-approval groups were formed by dividing at the mean of the particular experimental sample For the convenience of future researchers we report in the Appendix normative data on the populations studied in this volume and on several other groups

[4] Since subjects at the two universities did not differ on means or variances, we were able to analyze the data using the combined sample The analyses to be reported, however, were also carried out separately for the two groups, with results similar to those with the groups combined Thus, the findings that follow represent, in essence, the pooled results of a replicated experiment

TABLE 1 Differences between High- and Low-Need-for-Approval Groups in Expressed Attitudes

Question	High (N = 30) Mean	Low (N = 27) Mean	Diff	t
How enjoyable the task was	2 17	−0 70	2 87	3 53 *
How much they learned	5 37	3 22	2 15	2 63 †
Scientific importance	7 37	5 67	1 70	2 41 †
Participate in similar experiment	3 63	1 67	1 96	2 57 †

* $p <$ 01, two-tailed test
† $p <$ 02, two-tailed test

trast, on the first two questions the ratings of the lows are below the scale midpoints

A similar analysis was performed to assess the relationship of conformity to attitudes toward the experiment Scores on the Independence of Judgment scale were dichotomized at the overall mean (10 39) to yield high and low conformity groups The findings reported in Table 2 indicate only one significant difference between the mean ratings given by the two groups On question 2, the ratings of the high-conformity group as to how much they learned about their abilities and skills were significantly higher than the ratings of low conformers The questionnaire measure of conformity, then, was less efficacious in predicting favorable attitudes toward the spool-packing task than was the index of need for approval

Table 3 shows the results of the analysis of the Edwards *SD* scale Here, we again dichotomized scores at the mean (32 34) to obtain the high- and low-social-desirability groups Clearly, social desirability as measured by the Edwards scale is unrelated to the favorability of attitudes expressed toward the experiment

As a final step in the analysis of the data, scores on the *M-C*,

TABLE 2 DIFFERENCES BETWEEN HIGH- AND LOW-CONFORMITY GROUPS IN EXPRESSED ATTITUDES

Question	High (N = 31) Mean	Low (N = 26) Mean	Diff	t
How enjoyable the task was	1 31	0 39	0 92	1 04
How much they learned	5 27	3 58	1 69	2 02 *
Scientific importance	6 58	6 55	0 03	0 04
Participate in similar experiment	3 19	2 29	0 90	1 13

* $p <$ 05, two-tailed test

Edwards, and Barron scales were intercorrelated Scores on the *M-C* scale were once again significantly related to the scores on the Edwards *SD* scale ($r =$ 56), and the expected correlation with the Independence of Judgment scale was obtained ($r = -$ 54) The higher the score on the measure of need for approval, the lower the independence of judgment on the Barron scale Scores on the Edwards scale were uncorrelated with the inventory measure of conformity ($r = -$ 12)

TABLE 3 DIFFERENCES BETWEEN HIGH- AND LOW-EDWARDS *SD* GROUPS IN EXPRESSED ATTITUDES

Question	High (N = 14) Mean	Low (N = 15) Mean	Diff	t
How enjoyable the task was	0 36	0 47	−0 11	0 08
How much they learned	3 79	3 93	−0 14	0 12
Scientific importance	6 00	5 87	0.13	0 11
Participate in similar experiment	2 07	2 07	0	0

The results of the first experiment support a theoretical rationale which views socially desirable responding on personality inventories as the expression of a more general need for approval The uniformly more favorable attitudes of high-need-for-approval subjects may be accounted for in the following way The experimenters, as a result of their prestige and authoritative manner, reflected in their title, occupation, and behavioral aloofness, were perceived by high-need-for-approval subjects as persons whose favor was worth courting It seems not unlikely that to many college students professors are viewed as sources of approval gratification as well as dispensers of academic rewards and punishments In consequence, our high-need-for-approval subjects were strongly motivated to yield to the demands of the situation and to tell the experimenter that his experiment was interesting, important, personally informative, and worth returning to In contrast, subjects less approval dependent were better able to resist stating what seemed socially appropriate and offered, instead, more realistic appraisals of the experiment Presumably, the less favorable opinions of low-need-for-approval subjects reflect, in part, the greater freedom of this group from social pressures in the formation and expression of their beliefs The significant correlation between the *M-C* and Independence of Judgment scales would seem to support this formulation

The conformity scale, however, despite its significant relationship to the need for approval, did not serve reliably to discriminate the favorability of opinions expressed toward the spool-packing task Although, as we observed, there are certain apparent similarities between the two concepts, they do not have identical implications, and they do not equally well predict compliance with implied demand To restate an earlier point, conformity is better regarded as a mode of behavior determined along with other variables by dependence on the favorable evaluations of others

4

Three experiments in verbal conditioning[1]

Both socially desirable personality-test responses and compliance with the obvious wishes of the experimenter are neatly explicable by the concept of need for approval This explanation, however, is not crucial, and we might well develop several alternative and equally plausible accounts of these findings No evidence directly supporting the concept of *need* for approval has yet been presented, the goal-oriented character of the behavior of persons described as approval motivated requires a more critical demonstration

As usually defined in personality theory, reinforcement acts in the service of needs To put it the other way around, the stronger a given need, the more effective will be reinforcement appropriate to that need Certainly, then, given approving social reinforcement in a learning task, acquisition of the behavior to be learned should be enhanced for approval-motivated persons

For several reasons, the verbal-conditioning paradigm is particularly well suited to testing this proposition First, the verbal-conditioning situation has been widely considered as a miniature and simplified model for more complex interpersonal situations such as interviewing and psychotherapy (Krasner,

[1] This chapter considers two experiments previously published Experiment I is adapted from Crowne and Strickland (1961) Experiment II is drawn from Marlowe (1962)

The experimental task was then introduced with the request that the subject simply say words—all the words he could think of without using sentences or phrases and without counting The subject then proceeded to say words In the positive reinforcement condition, every plural noun he uttered was immediately followed by the experimenter's "Mm-hmm" and a head nod Subjects in the negative reinforcement condition elicited an "Uh-uh" for each plural noun We employed a nonreinforced control group to establish the base rate of plural nouns Subjects in this condition merely said words for the duration of the experimental time in the absence of any verbal reinforcement

The task was divided into five 5-minute periods during each of which the experimenter unobtrusively kept a frequency count of responses, tallying them as plural or nonplural At the end of 25 minutes the subject was stopped and asked the following questions

1 What do you think it was all about?
2 How did you go about deciding which words to say?
3 Did you notice any change in the kind of words you were saying?
4 Was there anything I did that you particularly noticed? (If "Mm-hmm" ["Uh-uh"] was not spontaneously mentioned, the subject was then asked, "What about my saying 'Mm-hmm' ['Uh-uh'] ?")

These questions were adapted from those employed in several studies reviewed by Krasner (1958) and were, of course, intended to serve as a test of awareness of the response-reinforcement link

In this experiment 145 introductory psychology students at Ohio State, 74 males and 71 females, served as subjects They were randomly assigned to the experimental or control conditions as they appeared for appointments For half these subjects the *M-C* scale was completed first, followed by the verbal conditioning task, for the remainder the order was reversed The *M-C* scale was not scored until after each subject had completed the experimental task There was thus no possibility that the experimenter's knowledge of the subject's approval score could influence the results

RESULTS

Figures 1 and 2, respectively, show the curves of the high- and low-need-for-approval groups under positive- and negative-reinforcement conditions and compare them with the curves of highs and lows in the nonreinforced control condition From the data plotted in Figure 1, it appears that the high-need-for-approval group shows the predicted increase in the frequency of plural nouns, while the plural-noun output of the low-need-for-approval group is intermediate between the highs and lows of

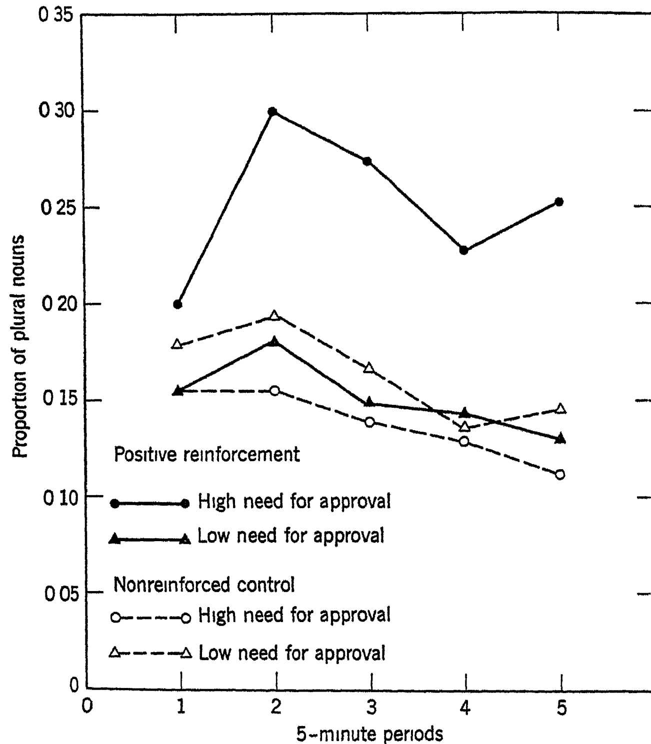


FIG 1 *Proportions of plural nouns given by high- and low-need-for-approval groups under positive-reinforcement and nonreinforced control conditions*

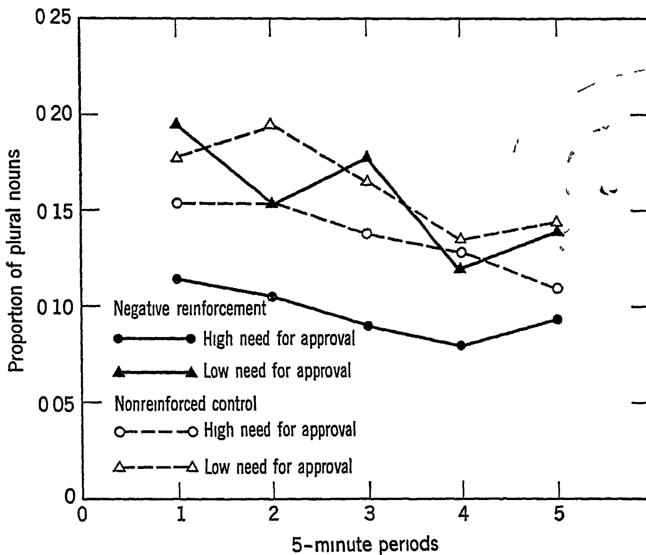


FIG 2 *Proportions of plural nouns given by high- and low-need-for-approval groups under negative-reinforcement and nonreinforced control conditions*

the control condition In Figure 2, the high group seems to decrease the use of plural nouns as predicted, with the low group again occupying a position between the nonreinforced high- and low-need-for-approval groups

Table 1 presents a condensed summary of the results of an analysis of variance (Grant, 1956) [2] There is a significant departure from a zero slope (Overall Trend $F = 3\ 83$, $p < 01$)

[2] Before the data were analyzed, 19 subjects were discarded—1 for counting, 15 from the negative-reinforcement condition, because they were aware of the conditioning procedure, 1 from the positive-reinforcement condition, for awareness, and 2 more randomly from the positive-reinforcement condition in order to obtain proportional cell N's for the analysis of variance The final experimental sample of 126 consisted of 42 subjects in each of the three conditions These subjects were further divided by dichotomizing the scores on the *M-C* scale at the overall mean (16 82) to yield the high- and low-need-for-approval groups

TABLE 1 Analysis of Variance of High- and Low-Need-for-Approval Groups under the Three Reinforcement Conditions

Source	*df*	*MS*	*F*
Overall Trend	(4)	031839	3 83 *
Linear	1	075927	6 09 †
Quadratic	1	006956	0 86
Cubic	1	043719	6 87 *
Quartic	1	000754	0 12
Between Group Means	(5)	261322	4 82 ‡
Need approval	1	003586	0 07
Reinforcement condition	2	284475	5 25 *
Interaction	2	367035	6 77 ‖
Between Individual Means	120	054212	
Between Individual Trends	(480)	008310	
Linear	120	012465	
Quadratic	120	008092	
Cubic	120	006363	
Quartic	120	006319	

*$p <$ 01
†$p <$ 02
‡$p <$ 001
‖$p <$ 005

It is clear that change in the proportion of plural nouns occurred in the course of the experiment Of major interest to the hypothesis are the significant Between Group Means effects Table 1 indicates that the reinforcers were influential ($F = 5\ 25$, $p <$ 01) and that they were differentially effective on the need-for-approval groups (Interaction $F = 6\ 77$, $p <$ 005) To isolate the differences within the positive-reinforcement condition, a t test was carried out between the means of the positively reinforced high- and low-need-for-approval groups for the first time period This analysis yielded a nonsignificant t of 1 21 We can conclude that at this early stage the verbal-conditioning

effect had not begun to separate the high-need-for-approval group from the other groups The differences between the reinforced high group and the low or control groups are to be found later in the task, in the consistent superiority of the reinforced high group during the remainder of the conditioning periods The smallest of these differences, that between the positively reinforced high- and low-need-for-approval groups for the fourth period, was tested by means of t The obtained value of 2 20 is significant beyond the 05 level

In the negative-reinforcement comparisons, t tests were also carried out between the means of the reinforced and nonreinforced groups Comparing the high- and low-need-for-approval groups under negative reinforcement, significant differences were obtained for the first, third, and fourth time periods However, "Uh-uh" was effective in differentiating the "punished" high-need-for-approval group from its nonreinforced counterpart only during the fourth time period The effect of verbal "punishment" was thus not as consistent in producing between-group differences as was the effect of approval A complete summary of the results is contained in Table 2, which shows the means and standard deviations of the high- and low-need-for-approval groups under the three conditions of the experiment

Finally, data bearing on an alternative explanation of the results remain to be examined It is not altogether impossible that the obtained differences in verbal conditioning are a function of intelligence Such a finding could occur if there were a relationship between intelligence and the *M-C* scale Scores on the Ohio State Psychological Examination (OSPE), a measure of verbal intelligence, were available on 34 of the positive-reinforcement subjects and 31 of the negative-reinforcement subjects OSPE scores did not differ between high- and low-need-for-approval groups in the positive-reinforcement condition ($t = 0\ 56$) or in the negative-reinforcement condition ($t = 0\ 23$) To assess any possible relationship between intelligence and verbal conditioning, correlations were run between OSPE scores and a total conditioning score obtained by summing over the time periods For the positive-reinforcement con-

TABLE 2 PLURAL NOUN RATIOS OF HIGH- AND LOW-NEED-FOR-APPROVAL GROUPS UNDER THE THREE REINFORCEMENT CONDITIONS

Group		Five-minute periods 1	2	3	4	5
Positive reinforcement						
High need for approval	Mean	197	300	272	229	256
	SD	116	202	223	143	199
Low need for approval	Mean	157	179	149	143	131
	SD	149	128	135	109	098
Negative reinforcement						
High need for approval	Mean	117	106	093	082	097
	SD	063	077	.098	069	134
Low need for approval	Mean	.194	155	176	128	143
	SD	146	.153	146	085	125
Nonreinforced control						
High need for approval	Mean	157	155	144	.134	117
	SD	099	115	.129	.101	097
Low need for approval	Mean	180	.194	.166	.138	148
	SD	120	138	114	.111	.181

dition $r = -06$, the corresponding r for the negative-reinforcement condition is -04 It is evident that intelligence, as measured by the OSPE, is unrelated to need for approval and verbal conditioning

Subjects whose need for approval was high tended to increase the relative frequency of the reinforced response class of plural nouns under positive reinforcement and tended to inhibit plurals when they were followed by punishment The low-need-for-approval groups failed to demonstrate consistent changes in rate of response under either positive or negative reinforcement compared with the base rate of plural nouns

The findings suggest that in the context of this experimental procedure a reinforcer connoting approval is more effective in increasing the rate of response than is punishment in inhibiting

response rate for subjects to whom approval satisfactions are important, yet disapproval was in general far more salient to subjects and for many was highly anxiety-arousing, as shown by their questions and comments, increased response latency, and other behavioral indications of disturbance. There were, however, no differences between high- and low-need-for-approval groups in these manifestations of anxiety, which were assessed by observational ratings during the conditioning period.

For all groups, there was considerable variability in conditioning. While the need for approval clearly seems to be one determinant of change in behavior in this experimental situation, it is undoubtedly not the only variable involved. Other studies of individual differences in "conditionability" or "responsivity" have predicted change in response rate from such diverse personality measures as manifest anxiety (Taffel, 1955), compliance in psychotherapy, test anxiety, and fearfulness in new situations (Sarason, 1958), achievement via independence (Krasner, Ullmann, Weiss, & Collins, 1960), and hypnotizability (Weiss, Ullmann, & Krasner, 1960). It is important to determine whether these and other personality variables operate in a meaningful constellation and whether the prediction of verbal conditioning (and other behaviors) can be increased by employing multiple measures.

We should note that no tautology is involved in this study in the finding that approval-motivated subjects increase their rate of response under conditions of approval-oriented reinforcement. subjects were carefully screened for awareness of the correct response, and only those who were unable to recognize the response-reinforcement sequence were retained for analysis. Thus, the effect of the need for approval was felt at a level of awareness below that which the subject was capable of verbalizing. In the main, subjects both high and low in need for approval tended to characterize the experimenter's "Mm-hmms" or "Uh-uhs" as generalized encouragers or discouragers or saw these reinforcers as pertaining to specific content categories which they were to pursue or abandon. There were no differences between high and low groups in this respect.

EXPERIMENT II

With a few exceptions (e g , Rogers, 1960, Salzinger & Pisoni, 1960), verbal-conditioning studies have seldom required ego-involved subjects to talk about personally relevant material to an actively interested listener Conceivably, it might be quite difficult to alter behavior which involves meaningful communication about oneself by such brief operant procedures Could one, for example, increase the rate at which subjects make positive self-references when the task is to describe one's personality? One study which bears directly on this question reported negative results (Rogers, 1960) Rogers was unable to demonstrate a significant increase over the operant level in the number of positive self-references emitted in quasi-therapy sessions Communications about oneself, since they are quite central verbal responses and are overdetermined by personality characteristics, are undoubtedly less susceptible to social influence attempts than responses of a more peripheral nature

Experiment II, then, sought to condition meaningful verbal behavior in an interview situation and to further test the need-for-approval–conditioning hypothesis

EXPERIMENTAL METHOD

The experimental task was a 15-minute individual interview conducted immediately after completion of the *M-C* scale[3] Each subject was told that the experimenter was making a study to learn what college students think and feel about themselves, and he was asked to describe his personality characteristics and traits If the subject remained silent for 20 seconds or more, the experimenter asked one of four questions, depending on what the subject had been talking about prior to that period of silence The questions dealt with the following areas relationships with people in general, relationships with family members, social life, and school work Every interview was tape recorded with

[3] To avoid bias, the need-for-approval score was not determined until after each interview was conducted and scored

the subject's knowledge and later transcribed, so that a content analysis of positive self-references and a reliability check could be carried out from the complete record

A total of 76 undergraduate students (34 males and 42 females) at the University of Kentucky served in Experiment II They were assigned at random to either the experimental or the control group For those in the experimental group, every positive self-reference was reinforced by the experimenter's "Mm-hmm," spoken in a flat monotone In the control group, the experimenter offered no reinforcement and restricted his responses to asking one of the four questions when a 20-second period of silence occurred

A positive self-reference was defined as any statement containing one or more personally favorable references (e g , "I get pretty good grades " "I was elected an officer in my fraternity ") A statement was defined as a sentence or complete "thought unit" after Dollard and Mowrer (1947)

Immediately after the interview, each subject was questioned to determine the extent of his awareness of the conditioning procedure The questions asked were similar to those in Experiment I

Reliability of the scoring of positive self-references was established by correlating the ratings of the experimenter with those independently obtained by another scorer who also worked from the typed transcripts [4] The correlation between the two raters was 89 based on 15 randomly selected interviews

RESULTS

The basic response unit used in the analysis was the ratio of positive self-reference (PSR) statements to total statements made The interview was divided into three 5-minute periods Mean PSR ratios for each 5-minute period and for the entire 15 minutes were computed for the high- and low-need-for-approval groups (formed by dichotomizing at the overall mean) in the

[4] These omitted the experimenter's reinforcements

TABLE 3 PSR RATIOS OF THE HIGH- AND LOW-NEED-FOR-APPROVAL GROUPS UNDER EXPERIMENTAL AND CONTROL CONDITIONS

Group		Periods			
		1	2	3	1–3
Experimental					
High need for approval	Mean	181	233	166	194
(N = 24)	SD	075	094	095	064
Low need for approval	Mean	178	130	115	140
(N = 14)	SD	086	079	058	062
Control					
High need for approval	Mean	229	121	124	155
(N = 18)	SD	117	082	083	063
Low need for approval	Mean	.203	114	075	129
(N = 20)	SD	082	067	019	045

experimental and control conditions These data can be seen in Table 3 The mean PSR ratios for the three periods are graphically illustrated in Figure 3

To test the hypothesis that under positive reinforcement high-need-for-approval subjects produce more positive self-references t tests of the differences in mean PSR ratios were computed For the first period no significant differences were obtained between the mean PSR ratios of any two groups During the second period the mean PSR ratio of the high-need-for-approval group is significantly greater ($p <$ 01) than the corresponding mean ratios of each of the other three groups For the last 5-minute period the PSR ratio of the high-need-for-approval group is significantly greater than that of the experimental lows ($p <$ 01), although not significantly different from the PSR ratio of the control highs When mean PSR ratios for the entire 15 minutes are compared, the high-need-for-approval subjects in the experimental condition are significantly ($p <$ 05) above each of the three comparison groups in the rate of positive

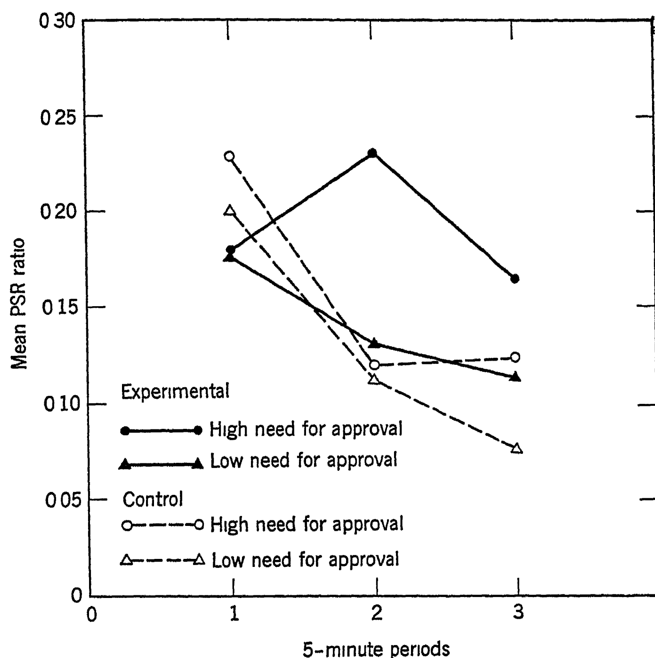


FIG 3 *PSR ratios of the high- and low-need-for-approval groups in the experimental and control conditions*

self-references Overall, then, paralleling Experiment I, individuals with a high need for approval produce significantly more reinforced responses under positive reinforcement than persons less strongly motivated for approval

That the differences between the high-need-for-approval subjects in the reinforcement condition and the three comparison groups are attributable to a conditioning effect and not merely to a decline in the PSR ratios of the other groups is indicated by the significant increase ($p < 05$) in the PSR ratio of the high-need-for-approval subjects from the first to the second period During the third period the high-need-for-approval subjects in the experimental group showed a significant decline ($p < 01$) in positive self-references

For the purposes of this study "awareness" was defined as an explicit statement by the subject that the experimenter said "Mm-hmm" each time the subject made a positive self-reference *None* of the experimental subjects was aware of the response-reinforcement sequence Of the 38 experimental subjects 27 reported that they noticed the "Mm-hmm" In general these 27 subjects stated that it signified "interest" on the experimenter's part or "encouragement" to continue talking It appears, therefore, that change in the rate of making positive self-references occurred in the absence of any awareness of the nature of the experimental manipulation

The results of Experiment II again indicate that differences in conditionability are in part attributable to individual differences in the need for approval, and they provide clear-cut confirmation of the first conditioning study. As in the first experiment, the behavior of high-need-for-approval subjects may be interpreted as reflecting greater sensitivity and responsiveness to social reinforcers than was shown by persons less strongly motivated to seek approval

It is very important to note that only a minor portion of the variance in responding with positive self-references can be attributed to the influence of the need for approval alone In the control group, which received no reinforcement, an insignificant correlation of 23 was obtained between *M-C* scores and PSR ratios for the entire 15 minutes For experimental subjects, a corresponding correlation of 42 ($p <$ 01) was obtained Thus, the increase in productivity for the reinforced high-need-for-approval group would seem to reflect their greater responsiveness to the reinforcements rather than any prior tendency to respond with positive self-references

A striking and significant ($p <$ 01) decline in the rate of making positive self-references occurred among the experimental highs during the third time period A similar decline occurred for the other three groups during the second period One might view this as a simple "exhaustion" phenomenon and conclude that the subjects just ran out of positive things to say about themselves An alternative and probably more likely explanation is that the decline is due to a reluctance on the

subjects' part to enumerate repeatedly their positive characteristics One subject in the experimental high-need-for-approval group, for example, spontaneously commented at the end of the interview, "It's awfully hard to sit here and toot your own horn"

Finally, it is worth noting that the results of this experiment add considerable generality to the frequently reported finding that verbal behavior is subject to control and modification through the use of operant-conditioning techniques Past studies have seldom required subjects to engage in personally meaningful communication, as a result, attempts to draw a parallel between verbal conditioning and communication processes in psychotherapy and interviewing have not been entirely convincing Self-reference communication can, however, be altered, particularly if reinforcements important to the subject are used Of course, the issue of whether other, more conflictual, kinds of self-references can be experimentally manipulated remains open

EXPERIMENT III [5]

In the two experiments we have just considered, and in most other verbal conditioning studies as well, subjects have been *directly reinforced* for critical responses Learning, however, does not necessarily require the immediate and direct reinforcement of overt behavior Miller and Dollard (1941) presented an extensive theoretical analysis of the role of imitation in social learning, and recently there have been investigations of behavior change under conditions in which a subject experiences vicarious reinforcement (Berger, 1961, Lewis & Duncan, 1958, Kanfer & Marston, 1963)

In vicarious reinforcement, a subject simply observes while the experimenter reinforces another person who is responding. During the observation period, the observer does not receive any reinforcement from the experimenter, nor does the observer make any overt response In the Kanfer and Marston study, sub-

[5] Portions of this experiment were reported at the Eastern Psychological Association meetings by Marlowe, Beecher, Cook, and Doob (1964)

jects who heard the experimenter say "Good" when a tape recorded subject made a critical response showed a significant conditioning effect during their acquisition trials

In this section we take up a study which attempted to relate the need for approval to vicarious reinforcement The finding that vicarious reinforcement can enhance conditioning is usually cast within a mediational framework (Lewis & Duncan, 1958). In effect, the subject is believed to respond covertly while he observes, if he makes a correct response, he may reinforce himself by saying "Good" or "I was right" Self-reinforcement during vicarious training, then, may lead to increased strength of the critical response class How may the self-reinforcement hypothesis be articulated with our findings regarding the approval motive? Is our previous emphasis on *social* approval inconsistent with the concepts of vicarious or self-reinforcement? Suppose these concepts were linked to approval-motivated behavior in a conditioning situation in the following way

Subjects undoubtedly differ in their responsiveness to the "demand characteristics" (Orne, 1962) of experimental situations Seldom are they merely passive participants; for many there is a strong desire to do a "good job" and to be a "good subject" It is true, of course, that other subjects are relatively indifferent to the evaluative consequences of their experimental behavior A basic theme which we have advanced about the behavior of approval-dependent persons is that they place a strong emphasis on correct behavior in order to please significant others Since high-need-for-approval persons are evaluation-oriented, we would expect them to be quite concerned with the meaning of experimental procedures In short, how can one do the "right thing" unless he has some idea of what the right thing is? Thus, we should find approval-dependent individuals more likely than low-need-for-approval persons to engage in self-reinforcement as part of the process of defining the purpose of the experiment and the adequacy of their own behavior

Self-reinforcement, to be sure, does not involve a significant other who dispenses indications of his approval for immediate overt behavior But it seems safe to assume that subjects under-

stand that we are interested in them and that anything they learn during the observation phase may be relevant to the final evaluation of their performance In this experiment, in fact, most subjects revealed during the postexperimental interview that they expected to do something after the observation period related to it

As a final step in detailing the rationale of this study and the predictions that were made, it is necessary to link the occurrence of vicarious reinforcement to the "evaluation-concern-self-reinforcement" process Our reasoning was straightforward Vicarious reinforcement can contribute to the process outlined above in at least two ways First, given the subject's desire to figure out what is important, vicarious reinforcement can provide a focus on the relevant response class Second, given that the subject self-reinforces his covert responses (or hypotheses), vicarious reinforcement can serve as an external confirmation that he is on the right track

We predicted, then, that high-, as compared to low-, need-for-approval subjects would be more responsive to vicarious reinforcement and would show a significant conditioning effect following an observation phase In an "Information" condition in which the subject observes the same behavior without the occurrence of reinforcement, no conditioning was expected, since the experimenter makes no response to orient the subject to the proper response class or to confirm the validity of the subject's self-reinforcements

EXPERIMENTAL METHOD

Prior to their appearance, the subjects—this time 45 Harvard freshmen paid $1 50 for their participation—were randomly assigned to one of three conditions the Vicarious-Reinforcement condition (*VR*), the Information condition (*I*), or a Control condition (*C*) Half filled out the *M-C* scale first and the others after conditioning The conditioning task was the Taffel (1955) technique, which requires the subject to make up a sentence using one of six pronouns and a verb The subject is handed a deck of unlined 3 x 5 index cards, each of which con-

tains a different verb in the simple past tense and the six pronouns The pronouns appear in a different order on each card

Subjects participated three at a time They were met by three male experimenters (undergraduates who said they were working for Dr Marlowe) and a fourth undergraduate who pretended to be a subject but who was actually a paid confederate Each subject was taken to a separate room by one of the experimenters, read the instructions, and handed an identical stack of 20 3 x 5 cards The experimenter simply listened and unobtrusively recorded the pronoun used on each trial as a measure of the subject's operant level of "I"-"We" usage The subjects then returned to the original room, where they met the confederate While two experimenters waited outside, the third experimenter conducted the observation phase of the study

This experimenter "randomly" selected one person (always the confederate), seated him opposite the three real subjects, and repeated the instructions The three subjects were told, "Observe what goes on Don't say or do anything, just observe" The confederate was given a stack of 40 cards which actually contained 40 sentences instead of just verbs and pronouns So, instead of making up sentences as the subjects believed he was doing, he read the 40 sentences Sentences starting with "I" or "We" occurred on 24 of the 40 trials (60 per cent against a chance occurrence of 33 per cent) The proportion of critical responses increased over trials and closely approximated a typical acquisition curve

In the *VR* condition the experimenter recorded the pronoun used and said "Good" when "I" or "We" occurred When the 40 trials were finished each subject immediately returned to the first room with his experimenter, was handed a stack of 40 cards, and asked to make up sentences again The three subjects received identical decks which contained no verbs in common with those used by the confederate or the first 20 operant trials The experimenter merely recorded the pronoun used on every trial

The procedure for the *I* condition was identical to that for the *VR* condition except that during the observation phase the experimenter only listened and recorded In the *C* condition the

observation phase was omitted After the initial 20 trials the subject sat doing nothing for 6–8 minutes (a period of time corresponding to the length of the observation phase) and was then handed the final stack of 40 cards

Subjects in the *VR* and *I* conditions were given an extended interview based on Spielberger, Levin, and Shepard's (1962) questions, slightly modified to fit in with the use of an observation phase The interview, consisting of about 15 questions, was intended to determine whether subjects were aware of the purpose of the experiment or the observation phase

The essential features of the design are summarized below·

	Part 1	Part 2	Part 3	Part 4
VR (N = 15)	20 operant trials	Observe confederate reinforced 40 trials	40 nonreinforced trials.	Interview
I (N = 15)	20 operant trials	Observe confederate (not reinforced). 40 trials	40 nonreinforced trials	Interview
C (N = 15)	20 operant trials	6–8 minute wait	40 nonreinforced trials	——

RESULTS

The unit of analysis was the mean number of "I" plus "We" responses per 20-trial block Table 4 contains a summary of the data in the form of means and standard deviations for high- and low-need-for-approval subjects in the three conditions[6] In

[6] The overall mean *M-C* score was used to divide subjects into the high- and low-need-for-approval groups The mean *M-C* scores obtained in the three conditions were virtually identical, as were the mean *M-C* scores of subjects who took the scale first and subjects who took the scale last

TABLE 4 "I"-"WE" USAGE BY TRIAL BLOCKS OF HIGH- AND LOW-NEED-FOR-APPROVAL GROUPS UNDER EXPERIMENTAL AND CONTROL CONDITIONS

Group		Trial Block 1	2	3
VR/high	Mean	7 50	9 25	9 63
(N = 8)	SD	1 94	3 07	1 49
VR/low	Mean	6 86	6 00	5 86
(N = 7)	SD	2 53	3 12	3 18
I/high	Mean	8 00	8 25	8 25
(N = 4)	SD	1 00	83	1 48
I/low	Mean	7 45	7 91	7 73
(N = 11)	SD	2 15	2 71	2 83
C/high	Mean	8 00	7 75	8 13
(N = 8)	SD	2 35	1 39	2 03
C/low	Mean	6 86	6 57	6 29
(N = 7)	SD	2 42	2 32	2 43

addition, data are presented for the combined high- and low-approval-need subjects ($N = 15$) in each condition in Figure 4

The conditioning findings in Table 4 are presented graphically in Figure 5 There is no evidence for the occurrence of a conditioning effect as a function of treatment condition alone The performance of the *VR* group is not significantly different than that of the *I* group, and neither differs significantly from the *C* group As can be readily observed from the flat slopes of these curves (Figure 4) neither the *VR* nor the *I* groups increased their rate of responding over the operant level

When the data are examined using the breakdown into high- and low-need-for-approval groups (Table 4 and Figure 5) there is evidence for the predicted conditioning effect for the high-need-for-approval subjects in the *VR* condition These eight subjects show a significant increase in mean "I"-"We" usage from operant level to the first block of 20 postobservation trials ($t = 2\ 60$, $p <\ 05$), an increment which is maintained during the final 20 trials ($p <\ 05$) No other group showed an increase

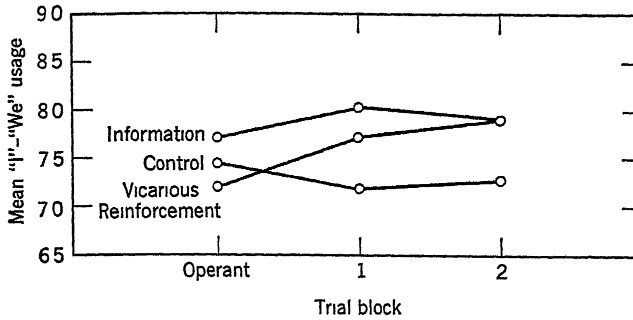


FIG 4 *Mean "I"-"We" usage for combined subjects under Vicarious-Reinforcement, Information, and Control conditions* (N = *15 per condition*)

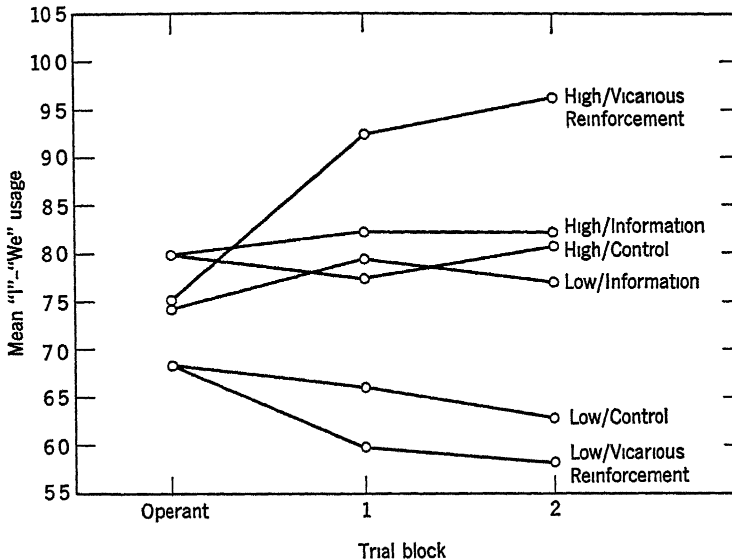


FIG 5 *Mean "I"-"We" usage for high- and low-need-for-approval groups under Vicarious Reinforcement, Information, and Control conditions*

in responses which even approaches conventional significance levels

The findings can also be expressed in the form of correlations For the 15 subjects in the *VR* condition the correlation between *M-C* scores and the increase in "I"-"We" usage from operant level to the first 20 postobservation trials is 45 ($p <$ 05) The corresponding correlation for the 15 subjects in the *I* condition is $-$ 19 (not significant) Thus, we find support for the prediction that high-need-for-approval subjects will show a significant verbal-conditioning effect under conditions of vicarious reinforcement The high-need-for-approval subjects in the *I* condition show no such effect The act of listening to the confederate, who produced exactly the same responses, does not lead to any conditioning in the absence of vicarious reinforcement

The "awareness" data obtained from the interviews were rated by three judges without knowledge of *M-C* or conditioning scores A 4-point scale was used to make these ratings The three judges made the same rating in over 90 per cent of the cases In no instance did a disagreeing judge differ from the other two judges by more than one scale point Six of the 15 subjects in the *VR* condition were rated as "aware" *The level of conditioning shown by these six subjects did not differ at all from that of the nine unaware* VR *subjects* In the *I* condition only one subject was rated as aware In the *VR* condition there was a slight, though statistically insignificant, tendency for low-need-for-approval subjects to report more awareness than approval-dependent subjects Thus, two out of eight highs received high awareness ratings, whereas four out of seven lows received such ratings

An incidental but interesting methodological finding is the fact that those subjects ($N = 20$) who completed the *M-C* scale first used significantly more "I"-"We" pronouns during the operant trials than those subjects ($N = 25$) who completed the scale last ($t = 2$ 78, $p <$ 01) This difference was not a function of the subject's *M-C* score, nor was the order of taking the scale related to the amount of conditioning The reason for this dif-

ference seems clear All the items in the *M-C* scale refer to the self and contain the pronoun "I" The effect of reading and answering the 33 items apparently induced a set to focus on oneself which carried over to the operant phase of the study and was reflected in greater "I"-"We" usage

A central assumption upon which the three experiments in this chapter were predicated is that approval-dependent persons participate in psychological experiments with a strong desire to do well and to be positively regarded Their evaluative dependence on an important other—the experimenter—leads them to actively seek approval and confirmation that they are performing adequately. The results of these experiments are consistent with this formulation only high-need-for-approval subjects showed a significant conditioning effect, and they did so only in those conditions where social reward was offered

The importance of some positive confirming response by the experimenter is highlighted by the findings of Experiment III High-need-for-approval subjects in the *I* condition (which lacked only the observation of social reinforcement) failed to show evidence of conditioning. There is no reason to believe that these subjects were any differently motivated than those in the *VR* condition Subjects in the *I* condition did not show an increase in response rate because the demand characteristics of the situation were not obvious, this condition lacked the cues defining a good performance

Verbal conditioning is obtained when in addition to the subject's positive motivation there occur positive and, by implication, evaluative responses on the part of the experimenter Such responses, even when they are vicarious, can serve to confirm for the subject that he has correctly diagnosed what is expected of him At an earlier stage in the observation phase, vicarious training can direct the subject's attention to the correct response class If the subject is indifferent to the evaluative implications of his behavior, or if he places an idiosyncratic interpretation on the meaning of the experimenter's confirming remarks ("Good"), then little conditioning is likely to occur For example, two subjects in the *VR* condition, though rated

fully aware, failed to demonstrate conditioning One subject remarked during the interview that it was "condescending" for the experimenter to say "Good", to the other subject it was "patronizing" Both obtained low need-for-approval scores Apparently, these subjects were so unconcerned with a favorable evaluation or approval that they simply did not interpret "Good" as a positive comment

For the third time we have found that awareness is unrelated to the level of conditioning manifested by high- and low-need-for-approval subjects Specifically, approval-motivated subjects, who are more conditionable, do not appear to be more aware We relied, of course, as have others, on postexperimental interviews and the subject's direct verbal report to determine the degree of awareness Such a procedure assumes that a subject is perfectly willing to reveal his understanding of the nature and the purpose of the experiment

Spielberger, Berger, and Howard (1963) attempted to replicate the approval-motive–verbal-conditioning results and found no relationship between the need for approval and conditioning. In their experiment, however, an undergraduate experimenter who apparently did not duplicate our efforts to establish the experimenter's importance to the subject administered the fairly simple and obvious Taffel (1955) procedure Another possible reason for the failure of this experiment may be that many subjects came to believe the experimenter was trying to influence them and resisted In such an explicit situation to conform by consciously giving the responses indicated by the experimenter would be to display one's susceptibility to social influence No one wants to be thought a conformer, whether he is in fact or not

Spielberger, Berger, and Howard's point, of course, is that awareness is the primary variable in verbal conditioning In effect, subjects do what the experimenter signals as correct, and the more clearly they understand what response he wants the more they show "conditioning" We have found no consistent evidence that our subjects were aware in this sense

It is incidentally worth noting that Spielberger et al found

no relationship between the approval motive and desire for social reinforcement (the experimenter's "Good") They see this as a predictive failure casting doubt on the validity of the need-for-approval scale To make such an interpretation, of course, is to assume that people are aware, and can tell someone else, of their evaluative dependence on others Personality questionnaires are constructed on the assumption that people often do not understand their own motives, nor can they describe them to others

One possibility in accounting for the fact that such a small number of our subjects demonstrated anything like an adequate definition of the response-reinforcement relationship is that our subjects, and perhaps particularly those high in the need for approval, were constrained not to tell us that they understood the purpose of the "Mm-hmm" or "Good," because to do so would "spoil" the experiment Our approval-dependent subjects might, in fact, have been quite aware but have failed to report it out of fear that the experimenter would be disappointed that his careful precautions to disguise the purpose of the experiment were found out Of course, if we managed to convey to our subjects a sincere interest in their hypotheses—an interest, moreover, associated with the experiment itself—this explanation goes out the window

Strickland (1962), who again replicated the need-for-approval–verbal-conditioning findings, rewarded her subjects for frank and thoughtful answers to awareness questioning in a postexperimental interview and found an incidence of awareness no greater than in a standard interview Dividing her subjects into aware and unaware groups on the basis of a strict criterion, she found greater conditioning in the approval-motivated aware group. Unaware high-need-for-approval subjects *still* gave evidence of verbal conditioning

A plausible alternative to the inhibition-of-report hypothesis is that the focus of approval-dependent individuals is on the evaluative consequences of their behavior rather than on testing hypotheses regarding the purpose of experiments Thus, the primary meaning of the experimenter's "Mm-hmms" or "Goods"

to approval-motivated subjects was a confirming or approving evaluation conveying, in effect, "You are doing well so far, I approve of you" The overriding question for our high-need-for-approval subjects, then, was, "How am I doing?" rather than "What is going on here?"

5

Conformity: group pressure and suggestibility[1]

The need for approval exerts an influence in situations far removed from the typical self-evaluative setting In the implied-demands study, we found strikingly compliant behavior among approval-motivated subjects Their favorable opinions toward spool packing were in sharp contrast to the more critical (and realistic) evaluations of subjects less motivated for approval We went on to suggest that social conformity may constitute a means of satisfying a need for approval from others In support of this proposition was the finding that scores on the *M-C* scale are significantly correlated with the Barron Independence of Judgment scale

But further, certain aspects of the behavior of conformists as described in other studies (Barron, 1953b, Crutchfield, 1955, Tuddenham, 1959) seem descriptive of people with a high need for approval If a person is concerned about the evaluative consequences of his behavior, it seems likely that yielding to the perceived demands of social situations becomes crucial for him Such a person may develop the expectancy that "fitting in" with others leads to personally desired goals and that minimal

[1] The first experiment described in this chapter is adapted from Strickland and Crowne (1962)

need satisfaction ensues from not conforming Conversely, from compliance and sensitivity to the evaluations of others may be inferred a disposition to seek approval and support from others A person so motivated is sensitive to the expectations of others regarding his behavior and seeks to respond according to their expectations as he perceives them

We must pause here to define more fully what we mean by conformity The simple correspondence of an individual's belief or opinion with the views of others does not by itself denote his conformity If he is, in fact, in agreement with the majority this is obviously not conformity, nor is conformity necessarily betokened by similarity in dress, manners, or other habits As Asch (1952) has pointed out, a meaningful definition of conformity requires that the individual demonstrably give up something of value or importance to him as the price of alignment with the group He must be shown to violate a personal norm, conviction, or strongly reinforced habit to be described as a conformer

In these experiments, then, conformity refers to the public statement of a judgment or opinion coincident with that of an incorrect majority in the absence of logical justification for that judgment In order to pose a conflict for subjects between a readily ascertainable objective fact and the opposing statements of a majority an "Asch-type" (Asch, 1951, 1952) conformity situation was employed We sought to determine whether the significant need-for-approval–conformity relationship previously found with the Barron scale would hold up in a behavioral situation where a price is exacted for conformity

EXPERIMENTAL METHOD

The first of the conformity situations used a procedure devised by Marlowe (1959) after one originally developed by Blake and Brehm (1954) to simulate the oppositional majority In this task the subjects were individually exposed to auditory stimuli which could be readily and accurately perceived Im-

mediately following each stimulus presentation were the responses of preinstructed accomplices of the experimenter, who gave inaccurate judgments on 12 critical trials out of a total of 18[2]

The auditory stimuli consisted of "knocks" recorded on tape by rapping on a table at a rate just fast enough to be credible as a test Following each series three accomplices, one at a time, announced their judgments of the number of knocks by calling out a number previously given to them by the experimenter The accomplices never disagreed in their judgments of the number of knocks

The subjects were 64 girls from introductory psychology classes at Ohio State They served individually in this experiment As each girl was brought to the experimental room she saw an impressive apparatus consisting of three tape recorders mounted in a vertical bank A set of earphones was connected to one of the three recorders, and a microphone appeared to be connected to a second machine Actually, this microphone served no purpose except to enhance the deception that the subject was really being recorded, and its cord was simply taped to the rear of the second recorder During the session the experimenter sat behind and to the side of the subject and recorded her answers The recording sheet was contained in a folder in which the recorded responses of the previous "subjects" were plainly in view The experimenter made certain that the subject could see this folder and the responses it contained during the presentation of the instructions, thus further enhancing the credibility of the experiment

The subject was instructed that she would hear on the tape a series of knocks followed by the answers of the three previous subjects, each of whom had stated how many knocks she heard

[2] Four graduate students were asked to listen to the tape recorded stimuli and to report correctly the actual number, disregarding the inaccurate judgments of the accomplices Their answers were in 85 per cent agreement with each other and with the number of knocks recorded Three of the judges had 92 per cent agreement

Her answers were to be announced into the microphone as soon as the third person had responded Each response of the naive subject was followed by a pause of about seven seconds, the subject then heard a second series of knocks and the answers of the three "previous subjects," and was then to announce her answer into the microphone The girls were told that their responses were being recorded on a second tape and that the purpose of the seven-second delay between trials was to leave recording room for subsequent subjects None was aware of the patent absurdity of this procedure—recording and playing simultaneously on two tapes

The experimenter read the instructions to the subject, concluding with this admonition

> Be sure to pay very close attention so you don't miss anything
> Do you have any questions? All right, now listen carefully and do as well as you can Performance on this task is highly related to mental alertness

The final statement was made in an effort to increase involvement in the task and to intensify the conflict between reporting the number of knocks actually heard and yielding to the distorted judgments of the "majority "

The girls all completed the *M-C* scale, approximately half of them before and half at the conclusion of the conformity task Following completion of the conformity situation each girl was interviewed to determine whether she should be excluded because she was aware of the deception Four subjects did state that they thought the experiment was "rigged," and they were eliminated from the analysis of the results Several subjects thought that the purpose was to "see if you would be influenced " These subjects were not dropped from the final sample since in response to the third question none indicated any suspicion In fact, many subjects who recognized that the purpose of the experiment involved "group" influence were among the more conforming This observation has also been noted by Asch (1952) and by Marlowe (1959), recognition of influence attempts does not necessarily inhibit conformity

RESULTS[3]

To test the hypothesis of greater conformity among high-need-for-approval subjects, scores on the *M-C* scale were dichotomized at the mean to form the high and low groups The significance of the difference between these groups in yielding to perceived group pressure was tested by means of *t* The results of this analysis are presented in Table 1 The difference is in the

TABLE 1 MEAN AMOUNT OF YIELDING OF THE HIGH- AND LOW-NEED-FOR-APPROVAL GROUPS

Group	*N*	Mean trials	Difference	*t*	*p*
High need for approval	30	6.73			
			2 63	3 08	01
Low need for approval	29	4 10			

predicted direction and highly significant, with the high-need-for-approval group behaving in a more conforming manner in the simulated group situation

We now need to look more closely at the tape situation as a measure of conformity If this experimental procedure is analogous to a group-pressure situation, the simulated majority should produce as much conformity as a present group In this experiment 24 out of 59 subjects, or 40 7 per cent, yielded 4 or less times on the 12 critical trials Thirteen subjects, or 22 0 per

[3] As a preliminary step in the analysis of the data the means and standard deviations were compared for those subjects who completed the *M-C* scale first and for the remaining half of the subjects, who initially participated in the behavioral situation No differences were obtained, and the data of these two groups were therefore pooled to yield a combined *N* of 59

cent, yielded on 9 or more trials These results approximate those of other studies in which the opposing majority was actually present (e g , Asch, 1952) In this respect, at least, the tape situation seems to be a faithful replica of the "live" conformity task

An additional possibility to be examined is that the tape situation is not a measure of conformity but constitutes a test of the ability to count auditory stimuli in a rapid sequence This did not appear to be true First, the accuracy of the judges in counting the number of knocks was quite high Also, Marlowe, employing the same task, found that nonconformists were more than 80 per cent correct on all 18 trials Clearly, it seems to have been possible for the subjects in this study to determine with a high degree of accuracy the actual number of knocks on any given trial It was further observed that the girls exhibited a considerable degree of conflict in the situation, for many it was a painful and anxiety-inducing experience Thus, the major determinant of behavior was not the ability of the subject to count rapidly presented stimuli but the degree to which she was affected by the simulation of group pressure

In formulating the hypothesis linking the need for approval with the behavioral variable of conformity it was formally recognized that behavior in a given situation can serve to satisfy many different needs Situations are categorized by the individual on the basis of implicit or explicit cues, by these cues he determines that certain behaviors will lead to desired goals or satisfactions and others to punishment and failure The cues in the conformity situation presented the subject with the conflict of having to decide whether "going along with the others" or "saying what I heard" was the behavior which would provide maximum satisfaction—or minimum "punishment "

The same subjects who endorsed socially desirable statements about themselves were those more likely to yield to group pressure This finding may have some suggestive implications for the problem of conformity We know little about the behavior of the conformer in situations other than those involving immediate group pressure In one study (Barron, 1953b), how-

ever, it was found that independents and yielders in an Asch situation differed in test-taking behavior, the yielders evaluating themselves as more "personally stable" and "well-adjusted" than the independents We found that approval-oriented individuals tend to achieve less pathological scores on the clinical scales of the MMPI than do persons with a weaker need for approval The image of the conformer that seems to emerge from this body of findings is of a person who not only is submissive to group pressure but who also appears to rely on cultural stereotypes in evaluating himself on personality tests Our results suggest that the behavior of the conformer serves his need for approval

At this point a note of caution is in order Although the experimental procedures and the behavior of subjects seem clearly to suggest that this experiment effectively simulated group pressure, we do not know what the effects might have been had the other group members actually been present How well would our conformity findings hold up when social pressure originates from persons who are present in the situation and in face-to-face contact with the naive subject? In the next experiment a conformity task with a "live" majority was used to obtain further evidence on the generality across situations of the relationship between the need for approval and conformity

METHOD

This conformity situation was an adaptation (*cf* Mangan, Quartermain, & Vaughn, 1960) of the Asch perceptual-discrimination procedure Briefly, the task consisted of a "discrimination problem" in which the subject was required on each of 20 trials to identify the larger of two clusters of dots The two groups of dots were presented tachistoscopically for one-second intervals by means of a 2 x 2 slide projector For the purpose of face validity each of the 20 slides differed in the geometric arrangement of the dots The larger of the two clusters was in each case clearly identifiable, as established by the 90 to 100 per cent accuracy achieved by 12 judges

Again, the subjects were recruited at The Ohio State University from introductory psychology classes Twenty-six girls volunteered for an experiment in "perceptual discrimination" and served as the naive subjects in the task The four confederates, two males and two females, were drawn from a pool of introductory psychology students and were specially trained in their roles Their task, of course, was to play the role of experimental subjects in as credible a fashion as possible

Each subject, together with the four accomplices, was met in the waiting room and escorted to a large room where she filled out the *M-C* scale To enhance our stooges' credibility, they pretended to fill out the questionnaire After completing the *M-C* scale the subject and the confederates were taken to the experimental room and given these instructions

> This is an experiment in perceptual speed I am going to show you some slides for a short exposure On these slides you will see two groups of dots labeled *A* and *B* I want you to tell me which of the two groups contains the larger number of dots

The order in which the subject and confederates responded was rotated on each trial so that each subject responded in each of the 5 positions on 4 occasions Four trials were not critical, i e , on these trials the confederates gave the correct response On 10 of the 16 critical or conformity trials the confederates gave incorrect answers without additional comment On the remaining 6 trials the first confederate answered correctly in a hesitant manner, whereupon he was immediately criticized by the remaining confederates who gave the wrong answer This confederate then changed his response and conformed to the incorrect majority These 6 trials are considered "strong pressure" trials They were included to assess the influence of amount of pressure on the disposition to yield We might expect that high-need-for-approval subjects would conform more than lows when exposed to strong pressure.

After completing the conformity task, each naive subject was carefully interviewed to determine the extent of her awareness of the purpose of the experiment

RESULTS

Our typical procedure was followed by dichotomizing scores on the *M-C* scale at the overall mean of the sample A subject's conformity score was the number of trials on which she agreed with the incorrect majority The mean number of conforming responses (without regard to trial category) for the high-need-for-approval group was 9 46, compared to a mean of 5 46 for the low-need-for-approval group The difference between these means is significant at the 02 level ($t = 2\ 56$) and supports the earlier finding of greater behavioral conformity among approval-motivated persons

To determine the effect on conformity of the amount of pressure exerted, the conformity scores were recalculated in the form of percentages to achieve comparability among three categories of trials These were the regular-pressure trials, in which the majority judgment was announced without any verbal exchange among the confederates, the strong-pressure trials described above, and the pressure trials immediately following the strong-pressure trials The last category was included to test a "carry-over" effect The percentage of conforming responses to these three groups of trials is shown in Table 2 Differences in the amount of yielding to these trial categories were assessed by the Wilcoxon signed-ranks test None of the differences between

TABLE 2 PERCENTAGE OF CONFORMING RESPONSES IN THE THREE GROUPS OF TRIALS

	Percentage of Conformity	
Type of Trial	High need for approval	Low need for approval
Pressure	56 41	35 90
Strong pressure	62 82	29 49
Immediately succeeding	57 69	38 46

trial groups was significant, nor was the approval motive related to conformity as a function of trial category

In the postexperimental interview, 10 of the 26 subjects voiced some suspicion as to the purpose of the experiment None of these 10 subjects, however, was aware that the procedure was "rigged"—that the confederates were not "real subjects" The questions raised by these subjects concerned the purpose of the experiment as it was described to them they rightly concluded that social influence was the variable of interest Of these 10 suspects, 7 were in the low-need-for-approval group By a Mann-Whitney *U* test, however, their yielding scores did not differ from the scores of the 6 unsuspecting lows The 3 high-need-for-approval subjects who questioned the purpose of the experiment conformed significantly less often than the remainder of the high group This finding, of course, works against the hypothesis

High-need-for-approval subjects have been shown to be significantly more conforming than lows in three quite different contexts a paper-and-pencil test situation intended to measure a "latent" disposition to conform, a simulated group-pressure situation, and a more natural group setting where the other persons were actually present In addition, the last two situations employed different perceptual stimuli as the material to be judged The results, overall, are highly confirmatory of the hypothesized relationship between individual differences in the strength of approval motivation and conformity to group pressure

The absence of any trial category effect in this experiment is due simply to the failure of the subjects to perceive the trial groups as differing in amount of pressure exerted The regular trials engendered so much pressure to begin with as to render ineffective our attempt to create a distinct increase in the amount of pressure exerted

Several more investigations which serve to extend the results of the conformity studies should be mentioned at this point. In all of these additional studies, direct group pressure was not employed. They may be characterized as investigations of "sug-

gestibility" or response to the implied "demand characteristics" of the experimental setting (Orne, 1962)

An extensive research program by Rosenthal (1963) and his students has investigated "experimenter bias" Their findings suggest that among relatively naive experimenters prior expectation—a directional hypothesis—is father to results To demonstrate biased experimental outcomes, experimenters are given directional hypotheses about the behavior of their subjects The subject's task is to rate photographs of people for "success" or "failure" Some experimenters are led to expect high, and others low, ratings *Every subject is given identical instructions* Rosenthal (1963) summarizes his findings as follows "In three studies, a total of 30 experimenters ran about 375 subjects In every one of these studies the lowest mean rating obtained by *any* experimenter expecting high ratings was higher than the highest mean rating obtained by *any* experimenter expecting low ratings from his subjects The three *p* levels were 004, 001, and 004" (p 270)

In several investigations, significant correlations were found between the need for approval and experimenter bias The average correlation over five samples was 74 High- and low-need-for-approval *subjects* were equally biasable One of the salient characteristics of the biasing experimenter is approval dependence His concern for a favorable evaluation and his desire to please lead to attempts to influence his subjects in the direction of the hypothesis given him He does this, presumably, by subtle means, little recognizing the effects of his wish to be a "good experimenter"

In a somewhat different area, Marlowe, Stifler, and Davis (1962)[4] sought to relate the need for approval to the classic phenomena of primary and secondary suggestibility With undergraduate males, approval motivation was not related to suggestibility as measured by the postural-sway test A strong relationship was found, however, between the need for approval and suggestibility as measured by the heat-illusion test This

[4] This was an undergraduate research project at Dartmouth College, carried out by Stifler and Davis under Marlowe's direction

technique asks subjects to indicate whether they detect heat in an element which, of course, never actually warms up. High-need-for-approval subjects reported heat significantly more often than lows ($p < .05$). The highs, then, were more suggestible, their chimerical report being influenced by the demand characteristics of the experiment.

To sum up, the experiments reported in this chapter have employed experimenters, settings, and tasks quite dissimilar in stimulus value and meaning for the subjects. The findings all converge on a single consistent conclusion: high-need-for-approval persons are more responsive than lows to perceived situational demands and are more likely to respond affirmatively to social influence. This is true whether the influence attempts are obvious or relatively subtle and whether they emanate from an individual or from a group.

6

Normatively anchored social behavior[1]

The point of departure for the experiments considered in this chapter is the apparent anchoring of the social behavior of approval-dependent persons in conventional, even stereotyped, cultural norms. We take up here two very different instances of social behavior: in the first experiment, language behavior, and in the second, goal setting in a situation in which a cultural norm is strongly implied.

Why choose such diverse instances of social behavior to demonstrate normative anchoring? Our intent was to test the generality of conventional, stereotyped behavior in approval-motivated persons and, hopefully, to show that in the absence of directly exerted influence dependence on social sanctions remains. Were this to be found, we would have evidence quite suggestive of a long and strongly reinforced history of social dependence and "other-directedness" in persons motivated for approval.

As a technique to measure conventionality in language in the first experiment, the word-association test was a clear first choice. The behavioral correlates of responses to the word-association test have been the subject of investigation since the procedure was first introduced by Galton (1879). The early emphasis on the utility of the test for lie detection and clinical

[1] The first of the two experiments in this chapter is from Horton, Marlowe, and Crowne (1963).

diagnosis (Jung & Riklin, 1904, Marston, 1920) has shifted in recent years to the search for relationships between word-association responses and personality traits (Jenkins, 1960, Sarason, 1959) Personality-oriented research in this area has focused on one particular aspect of free-association responses commonality of word associations, the tendency of an individual to give popular or common associates These efforts, however, have yielded meager results, and, as Jenkins (1960) noted, "Commonality has steadfastly refused to be linearly correlated with various scales of introversion, personality, achievement, and intelligence" Jenkins, however, did suggest that the degree of commonality manifested by a person is related to the extent to which he is like the "standard of the culture" on a variety of behavioral dimensions In a recent study, Jenkins (1959) had subjects take the word-association test under the standard "speed" instructions and a second time under a set to give "popular" responses High gain scores (an increase in common responses from "speed" to "popular" conditions) appeared to be related to "social sensitivity," according to reports of instructors who were acquainted with the subjects involved

A number of studies provide results which suggest that higher commonality scores may be obtained under "speed" than under "relaxed" instructions (Dunn, Bliss, & Siipola, 1958, Jenkins, 1959, Siipola, Walker, & Kolb, 1955) Jenkins (1960) has emphasized that commonality reflects strong intraverbal habit strengths and that personality variables are likely to influence commonality scores only under "relaxed" administration conditions where intraverbal habit strengths are less likely to be overly determinative

The hypothesis of this experiment was that high-need-for-approval subjects would obtain significantly higher commonality scores under "relaxed" instructions and under "popular" instructions than would low-need-for-approval subjects No difference between high- and low-need-for-approval subjects on commonality was expected under "speed" instructions where personality variables are less likely to be operative

EXPERIMENTAL METHOD

During class time, 295 introductory psychology students at the University of Kentucky completed the need-for-approval scale and the Kent-Rosanoff word-association test The association test was readministered to these subjects approximately four weeks later The subjects were divided into two groups for the testing Group I, which consisted of 153 subjects, took the word-association test first under "relaxed" and then under "popular" set instructions Group II, 142 subjects, took the association test under "speed" and then under "popular" set instructions A total of 259 subjects, 134 from Group I and 125 from Group II, were present for both test sessions The essential features of the instructions in each of the conditions were as follows

Relaxed Set

When you open these sheets you will see a list of 100 stimulus words After each word write the word that it makes you think of Work at your own speed, but do not skip any words Ready, go

Speed Set

Work rapidly until you have finished all 100 words When you are through, turn your paper over and write on the back the letter that appears on the board at that time Ready, go

After 4 minutes the letter A was printed on the blackboard Every 30 seconds thereafter a new letter was written down in alphabetical order.

Popular Set

I want you to take the same word-association test you took a few weeks ago This time you are to give popular responses That is, give the responses that you think would be given most often by young adults like yourself

Commonality scores were obtained for each subject on both the first and second word-association tests In addition, difference or gain scores between tests were computed for each subject present during both sessions Commonality scores were obtained for each subject by assigning a normative frequency of occur-

rence to each response given to the 100 stimulus words A second commonality score was obtained by assigning each subject one point for each response which was given by 50 per cent or more of the normative group [2] Thirty-three stimulus words elicited responses which were given by at least 50 per cent of the normative group Prior to analysis of the experimental variables, the relationship of the second commonality score—the 33-item commonality scale—to full-scale commonality scores was determined Correlations were computed between full-scale commonality scores and the number of primary responses given under the various instructional sets The average correlation was 97, indicating that commonality scores based on the 33 primary responses very adequately reflect the same word-association patterns as do full-scale commonality scores Accordingly, the following analyses were based on commonality scores derived from the 33 primaries

RESULTS

The findings on the influence of different instructional sets on commonality of word associations may be mentioned briefly Mean commonality scores under "speed" instructions are significantly higher than commonality scores under "relaxed" instructions ($p < 01$). As Jenkins (1959) also found, Group I and Group II subjects were able to increase their commonality scores significantly under "popular" set instructions ($p < 01$)

The major findings concerning the influence of the need for approval on commonality scores are presented in Table 1 For the entire sample, the mean need-for-approval score was 15 84 In this experiment, high-need-for-approval individuals were defined as those subjects with scores of 20 or higher on the *M-C* scale Low-need-for-approval subjects were those who scored 12 or below These cutting points very closely approximate the upper and lower 25 per cent of the sample The data reported in Table 1 provide partial confirmation of the hypothesis linking

[2] The normative data and the procedures for assigning commonality scores are fully described in Horton, Marlowe, and Crowne (1963)

TABLE 1 SIGNIFICANCE TESTS BETWEEN HIGH-NEED-FOR-APPROVAL (H) AND LOW-NEED-FOR-APPROVAL (L) SUBJECTS ON THE COMMONALITY MEASURE

Condition	N_H	N_L	Mean H	Mean L	t	p
Group I						
H vs L Relaxed	35	41	16.49	13.80	1.64	10*
H vs L Popular	31	33	20.61	17.79	1.38	—
H vs L Gain	31	33	3.94	3.64	0.18	—
Group II						
H vs L Speed	36	38	17.53	17.76	0.14	—
H vs L Popular	30	34	21.70	20.32	0.69	—
H vs L Gain	30	34	3.63	2.85	0.44	—

* Two-tailed test

the need for approval to commonality of word-association responses. High-need-for-approval subjects tended to obtain higher commonality scores than low-need-for-approval subjects ($t = 1.64$) in the "relaxed" condition (Group I). Under the "popular" set, this difference is in the expected direction but fails to approximate significance ($t = 1.38$). Contrary to prediction, gain scores (change in commonality scores from "relaxed" to "popular" instructions) are not related to need-for-approval scores ($t = 0.18$).

As expected, commonality under "speed" instructions and gain scores for Group II subjects are unrelated to need-for-approval scores. The failure of these subjects to show a significant effect under "popular" instructions was unexpected, although it may have been due to a carryover of responses from the previous "speed" situation. In any case, the effect is in the expected direction and is not grossly different from that found in Group I.

Two additional findings lend support to these conclusions. (1) In the "relaxed"-"popular" set group, those subjects who obtained commonality scores in the upper quartile on both the

first and the second administration of the association test scored significantly higher on the need-for-approval scale ($t = 2\ 50$, $p =\ 02$) than subjects who received commonality scores in the lowest quartile on both association tests (2) When commonality scores of high-need-for-approval and low-need-for-approval subjects were compared for the "relaxed" and "speed" conditions, high-need-for-approval subjects obtained scores of similar magnitude under both conditions, whereas low-need-for-approval subjects under "speed" instructions obtained significantly higher commonality scores than their counterparts under "relaxed" instructions ($t = 3\ 96$, $p < .01$) This finding indicates that high-need-for-approval subjects contribute little to the overall significant difference found between commonality scores under "speed" and "relaxed" instructions

Our data provide partial support for a relationship between commonality of word-association responses and the need for approval The greater number of common responses given by high-need-for-approval subjects probably reflects a combination of their defensiveness and reactivity to the language stereotypes of the culture It seems reasonable to assume that language habits, which are primarily the products of social learning, would be more consistent with cultural stereotypes in individuals with a high need for approval For persons strongly concerned with what the world thinks of them, language takes on a critical significance Not only does language have its customary importance as the vehicle for the communication of experiences, ideas, and feelings, but it is a primary means by which one is evaluated The person dependent on the approval of others must be able to use language to convey (even indirectly) his needs to significant others, and his use of language must serve, as well, a self-protective function That is, it becomes crucially important not to jeopardize the favorable evaluations of others in this as in other social acts

What we appear to observe, then, is the tendency among approval-motivated individuals to restrict their associations to the conventional, popular, or common It seems a plausible hunch that the character of associations in the word-association test might well reflect the quality of verbal communication

in the larger domain of "real life," particularly in those situations in which social evaluation is salient To sum up, "Watch what you say!" would seem to be the self-imposed injunction of individuals concerned with the approval of others

The extent to which the need for approval is associated with cultural stereotypes in language is apparently not related to the amount of increase in commonality scores from "relaxed" or "speed" instructions to "popular" instructions Our findings here are contrary to Jenkins' (1959) suggestion that gain scores might reflect a heightened "social sensitivity" But it should be noted that the effect, if any, would be difficult to detect because of the initial higher commonality scores of our high-need-for-approval subjects A more sensitive test of Jenkins' suggestion would be afforded by a comparison of high- and low-need-for-approval subjects low on initial commonality scores Such a comparison was not possible in this study, because the number of subjects meeting the initial qualification was small

A result that appears to have important implications for research on word association is that need-for-approval scores were not predictable from commonality scores Although the need for approval was related to commonality scores under "relaxed" conditions, the reverse was not true with the entire range of commonality scores Methodologically, this effect suggests that commonality may be such a complex or multidimensional variable that it would be difficult to determine its relationship with other psychological characteristics if predictions were made only on the basis of commonality scores

This experiment raises another issue for research in the area of word association Considering the effects of instructional variation, it is clear that individuals give significantly fewer common responses when the association test is administered under "relaxed" as compared to "speed" instructions This finding is entirely consistent with Jenkins' (1960) suggestion that a reduction in the number of common associations can be expected under testing instructions that provide greater opportunities for personality variables to operate Under pressure to respond quickly, overlearned, stereotyped language habits are likely to overshadow any influence that personality variables might exert

This finding is of considerable significance for the use of the word-association technique as it is customarily employed in diagnostic settings. As typically used, the test is administered under a set to "respond as quickly as possible." Since idiosyncratic responses are those with the greatest diagnostic importance, the usual "speed" set would seem to reduce the effectiveness of the technique in eliciting diagnostically meaningful material. It appears that fewer stereotyped (nondiagnostic) responses are likely if the word-association test is administered under relaxed conditions.

Our approval-dependent subjects did, however, tend to produce more common responses even under the most optimal conditions for the appearance of personally meaningful responses. In a somewhat different form we shall return in Chapter 11 to the problem of constricted and stereotyped responding—this time on projective tests.

If approval-motivated persons tend to produce verbal associations anchored in conventional usage, would their behavior in other situations similarly reflect responsiveness to perceived social norms? Barthel (1961) used a goal-setting situation to further test the cultural-anchoring hypothesis.

Most of the research on goal-setting behavior has focused on the effects of success and failure (Hoppe, 1930), assessed the defensive implications of behavior designed to avert anticipated failure (Rotter, 1943, Sears, 1940), or attempted to determine the relation of goal setting to such personality variables as nAchievement (Atkinson, Bastian, Earl, & Litwin, 1960). Goal-setting behavior, or level of aspiration, may be regarded, however, not only as individually diagnostic (for example, in the kinds of avoidant defenses people use) but in certain situations may be influenced by perceived cultural norms as well.

Consider the normative expectations in, say, a dart-throwing game where one is free to choose his own distance from the target. To throw from a point very close to the target would, of course, make the achievement of a high score much easier, so easy, in fact, that one would open himself to derision for setting his goal at such a low level of risk. On the other hand, there are norms regulating great degrees of risk as well. Should

an individual stand at a great distance from the target his level of aspiration becomes tantamount to boastfulness and is likely to be treated by others as such Thus, two culturally originating forces operate to restrict the range of choice and, for the individual dependent on the favorable evaluations of others, to make it probable that his level of aspiration will be limited to a range of intermediate risk Conscious of himself and concerned with obtaining the positive regard of others, the approval-dependent person is more likely to set his goals in closer conformity to the norms of the culture as he perceives them

In this experiment the normative anchoring of goals was defined by the variability of a subject's choices of distance from the target Concern about the evaluations of others and reliance on external standards should lead to the restriction of behavioral choices to those most culturally acceptable In the dart-throwing task of this experiment the subject has no opportunity to compare his performance directly with others, and he must choose his own level of difficulty on the basis of prior knowledge of social norms and his own expectation of success If we can assume equal expectations of success in tasks of this kind, then the effect of approval dependence should be to restrict the individual to a limited range of choices—that is, limited variation around his own mean

METHOD

Forty undergraduate males took part in the experiment For the dart-throwing task distances were marked off on the floor in 1-foot intervals from 1 to 20 feet Barthel gave the following instructions to each subject

Today you are going to be playing a game of darts Before we get into the main part of the game, I would like you to have the opportunity to take some practice shots from specified distances along the lines in front of you To begin, take one practice shot from the 9-foot line Now try one from the 14-foot line, now the 11-foot, now the 2-foot, and finally, try a shot from the 6-foot line

Now we can get into the main part of the game to see how good you are You will be given a number of trials, each trial will consist of 5 throws of the darts Now for any of these trials you can stand at any of

these lines that you like But once you have decided where you want to stand for any one trial, remain in that position until those 5 trial shots are over Then you may decide what you would like to do for the next trial You could either remain where you were or move to another position

The idea of the game, of course, is to get as high a score as possible Your score for each trial will be the points that you make on the dart board, multiplied by the distance at which you are standing for that trial For instance, if you were to get 10 points on the dart board and had been standing at the 1-foot line, your total score would be 10 If you had been standing at the 10-foot line, your total score would have been 100, and if you had been standing at the 20-foot line, your total score would have been 200 Do you have any questions? Okay, after each trial, I will record your score and then you can pick up the darts and move on to the next trial

Subjects were then given 20 trials of 5 throws each The index of the dependent variable—the range in which goals were set—was the variance of each subject's choices

RESULTS

Subjects were divided into high- and low-need-for-approval groups by breaking the *M-C* scale distribution at the overall mean, and the goal-setting variances of the two groups were compared by means of *t* Table 2 shows these data and the results of this analysis There was a highly significant tendency for approval-dependent subjects to restrict the range of their choices of distance from the target However, we must ask whether differences in success were related to selection of dis-

TABLE 2 Mean Variability of the High- and Low-Need-for-Approval Groups

Group	*N*	*M*	*SD*	*t*	*p*
High need for approval	19	7 23	5 19	3 18	005
Low need for approval	22	13 44	6 93		

tance from the target Barthel obtained a performance score for each subject and then computed the partial correlation between the need for approval and the variability of subjects' choices with success controlled This correlation was – 32 ($p <$ 05) the higher the score on the *M-C* scale, the more attenuated the range of goal setting For comparison purposes, the correlation between the need for approval and the variance of a subject's choices was – 34 Clearly, then, choice of distance from the target was uninfluenced by differential success or failure

In explaining these findings a basic fact to be recognized is that in American culture there is an implicit norm that one should neither make a task too easy nor show off by making it too difficult High-need-for-approval subjects seem to be more influenced by these implicit cultural standards of goal setting, among approval-dependent individuals, level of aspiration is, as a result, more restricted and cautious

Barthel (1963) repeated this experiment, using a measure of generalized expectancy of success and administering the dart-throwing task under neutral, self-esteem-threatening, and positive-self-esteem conditions He hypothesized that approval-motivated subjects, especially those with a low generalized expectancy of success, would exhibit greater defensive rigidity in goal setting than subjects less dependent on approval or approval-motivated subjects with a high expectancy of success The measure of defensive rigidity was the variance in choices of distance from the target Thus, Barthel interpreted behavioral restriction in this situation as an indication of defensiveness—specifically, rigidity

The predictions were confirmed approval-dependent individuals, especially those with a low expectancy of success, exhibited greater rigidity in the neutral condition than did less approval-dependent subjects or high-need-for-approval subjects with a high generalized expectancy In this condition, the least rigid subjects were those low in the need for approval with a high success expectancy The goal-setting rigidity of high-need-for-approval–low-expectancy subjects increased under threat to self-esteem, and a marginally significant trend was noted for this

group to show less rigid behavior when self-esteem was experimentally enhanced

Barthel suggested that these findings on approval dependence

> are reminiscent of the maladaptive results of the "ideal self" described by Horney [1950] As Horney has pointed out, the development of an unrealistic self-image often results in attempts on the part of an individual to live up to the idealized picture through (1) overdependence on others, (2) the constant search for means to bolster the self-image, and (3) fear of making mistakes and extreme sensitivity to criticism This individual attempts to ward off disconfirmation of his ideal by covering up personal flaws before others become aware of them Such a person is very similar to the individual who places a high value upon approval but holds a relatively low expectancy for success (p 120)

Barthel's findings imply that defensiveness contributes to the normative anchoring of the behavior of approval-dependent individuals Part III of this book contains a series of experiments designed to test the implications of the "defensiveness hypothesis" The fact of normative anchoring, however, seems clear in two very different situations, individuals motivated to obtain the approval of others appear to rely on perceived cultural norms as behavioral models

7

Task categorization and perceptual defense[1]

The greater dependence of approval-motivated individuals on social sanction for their behavior is associated with sensitivity to interpersonal cues in social situations—cues which mediate expectancies for those behaviors which will insure approval or yield disapprobation and punishment As a general statement this is undoubtedly true, and the verbal-conditioning experiments give empirical support There are, however, some important and untested implications for example, how would approval-dependent persons deal with the socially disapproved—with, say, obscenity? In an experimental setting in which the subject is asked to recognize and report obscene material, would concern for approval lead to inhibition or would the result be compliance with the apparent demands of the situation by an open and frank acknowledgement? This is the problem of the experiment we now take up

The procedure chosen for this experiment was the "dirty word" perceptual-defense paradigm first introduced by McGinnies (1949) The perceptual-defense situation can be seen as posing a conflict for the subject between the demands of the task for quick and accurate recognition versus a typically long history of inhibition associated with socially unacceptable

[1] This chapter is a revised version of Barthel and Crowne (1962)

words It thus affords an excellent opportunity to study the effects on conflict resolution of motive strength and the manner in which situational cues are interpreted by the subject

The cues in this experimental situation are likely to lead the subject to arrive at one or the other of two conclusions A subject might conclude, on the basis of the stress on speed and accuracy in the instructions, the tachistoscopic presentation of stimuli, the general set to "do as well as you can" in psychological experiments, and perhaps his own relative lack of inhibition in this area, that rapid recognition is essential for favorable evaluation of his performance Alternatively, discovery of the scatological nature of some of the words might lead the subject to wonder about the experimenter's interest in probing his reactions to such material The latter interpretation is reminiscent of Howes and Solomon's (1950) account of the plight of the naive college student confronted with the awesome realization that he may have seen an obscene word and must, to obey the instructions, report it to an omniscient experimenter who might then draw personal conclusions about him Thus, two salient elements in this experimental situation appear to be equally capable of influencing, in mutually exclusive ways, the orientation of the subject toward the task It might well be assumed that these situational categorizations have their antecedents in the child-rearing backgrounds of subjects—perhaps in parental acceptance or punishment of childhood explorations of sexuality and obscenity

How the subject categorizes the task assumes an even greater importance when the need for approval is taken into account Because of his concern for a favorable evaluation the approval-oriented individual is likely to be more affected by his interpretation of the purpose of the experiment and the criteria for the evaluation of his behavior It follows, then, that a high-need-for-approval subject who categorizes the task on the basis of its socially disapproved aspect (the taboo words) should display the greatest discrepancy between the recognition thresholds for neutral and taboo words At the other extreme, the approval-motivated subject for whom rapid recognition and

achievement cues are pre-eminent should achieve the smallest discrepancy between taboo- and neutral-word recognition thresholds. Low-need-for-approval subjects should be less influenced by the conflict between task requirements and the inhibition associated with taboo words. Intermediate discrepancy scores would be expected from this group.

Before taking up the experiment itself we need to consider briefly the nature of the perceptual-defense phenomenon. A persistently thorny issue in perceptual-defense research lies in the failure of the typical experimental procedures to converge on an unambiguous interpretation of the underlying processes (Eriksen, 1951; Howes & Solomon, 1950; Postman, Bronson, & Gropper, 1953). After more than a decade of experimentation it is still not clear whether failure to report the perception of threatening symbols or objects involves an increment in the recognition thresholds for such stimuli or is simply a function of the withholding of socially taboo responses. This experiment was specifically designed to assess the effects of motive strength and expectancies regarding the evaluative consequences of task behavior. However, the experimental procedure required subjects to record their responses in writing, thus permitting an exact scrutiny of the kinds of omissions and distortions made by subjects. A failure to make any entry at all might be expected in the case of perceptual blocking or defensive distortion. Suppose, on the other hand, that erasures or obliterations of taboo responses prior to the final and correct recording were found. This would establish a strong case for response withholding. Erasing or scratching out a recorded entry on a taboo-word trial would seem likely to entail the inhibition of a response whose nature has already been recognized and not the prior interference with "perception" by an unconscious defensive process.

EXPERIMENTAL METHOD

There were, as usual, two parts to the experiment: the behavioral measure (perceptual defense) and the *M-C* scale. In a pilot study employing a slightly different procedure from the

one finally adopted neither the scores on the personality inventory nor the reporting of taboo words was found to be affected by order of administration For the sake of convenience the need-for-approval scale was thereafter routinely administered after the experimental task

The measure of perceptual defense was a modification of McGinnies' method Ten words were used in the experiment, four taboo and six neutral words in the following order "apple," "dance," "ounce," "whore," "alert," "penis," "prize," "bitch," "yeast," "screw "

Following the practice trials on which the first two neutral words were presented, the complete experimental series of eight words was shown tachistoscopically, first at an exposure of 0 01 second and then, in the remaining trials, at exposures successively 0 01 second longer for each trial. At the end of the tenth trial all words that had been correctly reported were eliminated, and thereafter elimination of recognized words took place at five-trial intervals Correct recognition consisted of the exact recording in writing of the taboo and neutral words, and the task was continued to the trial at which this criterion was met for all words

The perceptual defense score was the mean difference between the number of trials required for recognition of the taboo words and the number necessary to recognize the neutral words As supplementary indices, frequency counts were made of erasures or obliterations of taboo words prior to correct reporting and of the employment of words similar in structure to the taboo words Structurally similar words were incorrect substitutes for taboo words containing at least two correctly juxtaposed letters, e g , "whose" for "whore" and "penal" for "penis." In the case of erasures and obliterations it was frequently possible to determine that what had been effaced was the correct taboo word

In this experiment, 129 introductory psychology students, all girls, served one at a time The selection of females was deliberate the encounter between a female subject and a male experimenter would, we hoped, intensify the conflict over reporting taboo words

When brought into the experimental room the subject was seated facing the tachistoscope aperture and introduced to the task The important parts of the instructions were as follows

In this experiment you are going to identify some words which will be shown to you through this tachistoscope window As you can see, there is nothing there but a blank screen with a small square in the center Each word that you are to identify will appear for a very brief interval in the center of the small square Write what you see in the correct blank If you are not sure of what you see, it's OK to guess If you see nothing, put a dash in the blank

OK, let's try a couple of words for practice

At the conclusion of the task and after finishing the *M-C* scale, each subject was asked the following questions

1 What was the experiment all about?
2 Did you feel that the type of words that were used had anything to do with it?
3 Did you react differently to some words as compared to others?

Responses to the first question were used to determine the manner in which the subject categorized the task and to provide a basis for inferring the subject's beliefs about the criteria for evaluation of her behavior Although differing widely in their accounts of the experiment, the subjects' answers were generally amenable to classification into two categories (1) responses naming perceptual keenness as the focus of the experiment ("perceptual speed" classification), and (2) answers referring to the social disapproval associated with the reporting of taboo words ("social disapproval" classification) [2] Six judges, both staff and graduate students in clinical psychology, were in 98 per cent agreement in placing these responses into the

[2] The following are illustrative examples of responses placed in the two categories

"Perceptual" categorization "To see how fast you can see words, I guess" "To determine the rate of seeing something fast Your rate of reading" "Disapproval" categorization "I don't know—to see how fast you see different words, how well you can perceive bad words" "I wasn't sure because I didn't believe it—seeing the vulgar words"

two categories[3] None of the judges had any knowledge of the subjects' need-for-approval scores or their performance on the experimental task The judges' ratings were made on the original interview protocols recorded at the time the interview was conducted and prior to the experimenter's scoring of the personality inventory or the experimental data sheets

In the assignment of subjects to the two categorization groups, approximately equal numbers were placed in each classification

RESULTS

The experimental design yielded four groups of subjects from the combinations of high or low need for approval and "perceptual speed" or "social disapproval" task categorization A two-by-two analysis of variance was performed to test the effects of the independent variables The two independent variables were found to be unrelated, with equal numbers of high- and low-need-for-approval subjects falling in each task categorization[4]

Table 1 presents the analysis of variance for the major hypothesis The F ratios for both need for approval and task

---

[3] Eight subjects were excluded prior to the ratings because of their failure to provide meaningful interview responses (e g, "I don't know" "I wasn't sure") Elaborations could not be obtained from these subjects The responses of an additional six subjects were too vague and task irrelevant to be reliably rated on the criteria of task categorization Typical of these responses were "Which words were first and which were last" "Personality" and "Individual reactions" These subjects were also dropped from the final data analysis However, had the latter six subjects been assigned to task categorization groups on the basis of the judges' unreliable classifications of their responses, the results would not have been altered

[4] In addition to the 14 discarded cases described above, 19 other subjects were excluded One of these could not be interviewed after the task due to time pressures Another 18 subjects were dropped chronologically in order to obtain equal cell N's for the analysis of variance The perceptual-defense scores of these 18 subjects did not differ significantly from the mean scores obtained in the groups from which they were eliminated, and the same trends were observed in the group of discards as were found in the final experimental sample

TABLE 1 Analysis of Variance of the Effects of Need for Approval and Task Categorization on Perceptual Defense

Source	*df*	*MS*	*F*
Need for approval	1	365 49	17 81 *
Task categorization	1	418 75	20 41 *
Interaction	1	222 05	10 82 †
Within (error)	92	20 52	—
Total	95	—	—

* $p <$ 001
† $p <$ 005

categorization attained significance at beyond the 001 level. However, the highly significant interaction of the independent variables warranted further analysis, as presented in Table 2 The difference between high-need-for-approval subjects offering "perceptual" as opposed to "disapproval" categorizations was highly significant, with the "social disapproval" group showing higher recognition thresholds Differing interpretations

TABLE 2 Mean Differences in Taboo-Neutral-Word Recognition Thresholds between the Experimental Groups

	Task Categorization						
	Perceptual			Disapproval			
	N	*M*	*SD*	*N*	*M*	*SD*	*t*
High need for approval	24	14 31	2 62	24	22 34	6 19	5 82 *
Low need for approval	24	14 26	4 17	24	14 58	3 68	0 28
t		0 05			5 28 *		

* $p <$ 001, two-tailed test.

of the purpose of the experiment by high-need-for-approval subjects affected to a striking degree their performance on the perceptual-defense task The mode of task categorization did not influence the recognition thresholds of low-need-for-approval subjects, and the magnitude of perceptual-defense scores of low-need-for-approval subjects employing a "socially disapproved" categorization did not differ from the performance of high-need-for-approval subjects using a "perceptual speed" classification Finally, it is clear that the significant need-for-approval main effect was contributed by the high-need-for-approval–"social disapproval" group

To investigate further the task behavior of the four groups, the usage of similarly structured words and the incidence of erasures were analyzed The results of *t* tests on the use of similarly structured alternatives to taboo words are presented in Table 3 These findings show that approval-motivated subjects who perceived the task as involving social disapproval employed significantly more words similar in structure to the taboo words prior to correct reporting than did high-need-for-approval–"perceptual speed" classifiers or the low-need-for-approval–"social disapproval" group

Table 4 presents the mean number of erasures for the various

TABLE 3 MEAN DIFFERENCES IN USAGE OF STRUCTURALLY SIMILAR WORDS BETWEEN THE EXPERIMENTAL GROUPS

	Task Categorization						
	Perceptual			Disapproval			
	N	*M*	*SD*	*N*	*M*	*SD*	*t*
High need for approval	24	9 92	5 45	24	25.54	12.95	5 44 *
Low need for approval	24	12 67	11.80	24	13 12	7 57	0 16
t		1 04			4 06 *		

* $p <$ 001, two-tailed test.

TABLE 4 MEAN DIFFERENCES IN ERASURES BETWEEN THE EXPERIMENTAL GROUPS

	Task Categorization						
	Perceptual			Disapproval			
	N	*M*	*SD*	*N*	*M*	*SD*	*t*
High need for approval	24	0 21	0 51	24	1 12	1 48	2 84 *
Low need for approval	24	0 37	0 57	24	0 62	0 92	1 14
t		1 00			1 40		

* $p < 01$, two-tailed test

groups The pattern of these findings approximates the results of the preceding analyses The mean number of erasures of taboo words prior to correct reporting was greatest among approval-oriented "social disapproval" classifiers and least among high-need-for-approval "perceptual" classifiers Low-need-for-approval subjects were intermediate Those approval-motivated subjects for whom the taboo nature of the words was a salient aspect of the experiment appear to have delayed their report by the avoidant means of alternative hypotheses and erasures or obliterations It is interesting to note that similarly structured hypotheses generally tended to precede the occurrence of erasures or obliterations, suggesting a process of increasing awareness of the taboo words and an increasingly conscious and deliberate attempt to avoid the embarrassment of having to acknowledge them Lending support to this observation is the fact that very few erasures or obliterations of neutral words occurred, erasures thus assume the character of a defensive or avoidant maneuver

As a means of additionally clarifying the issue of response withholding versus perceptual defense, the answers to the three interview questions were examined for the specific admission of delay of taboo-word responses The frequency of such admis-

sions was not significantly related to the need for approval The results of a χ^2 test of the degree of association of task categorization and admission of response withholding are presented in Table 5 A significantly greater number of subjects employing the "social disapproval" categorization admitted response withholding than did subjects categorizing the task as a measure of perceptual speed It is important to consider that the influence of task categorization was found in high-, not low-, need-for-approval subjects, despite the lack of relationship between the need for approval and admission of response withholding

At this point, a more careful examination of the taboo-neutral-word discrepancy scores of the two low-need-for-approval groups and the approval-motivated "perceptual speed" group was undertaken to determine if all the variance in perceptual defense could be accounted for by response withholding These groups were chosen because they had the smallest discrepancy scores and least frequently employed the avoidant measures of alternative hypotheses and erasures Tested against the hypothesis of a zero difference between the number of trials to recognize taboo versus neutral words, each of these groups produced highly significant t ratios Thus, a residual discrepancy existed for which the analyses failed to provide an explanation In an attempt to account for this residual, we utilized the data from the pilot study, in which the taboo and neutral words were each individually presented until correct recognition occurred A plausible alternative to the perceptual-defense hypothesis in

TABLE 5 FREQUENCY OF ADMISSION OF RESPONSE WITHHOLDING

	Task Categorization	
Response Withholding	Perceptual	Disapproval
Admission	3	27
Failure to admit	45	21

$\chi^2 = 25\,65,\ p < \,001$

accounting for the unexplained remainder might be a "shock" hypothesis Taken by surprise, subjects tend to delay their report of having seen socially unacceptable words until they have arrived at an interpretation of the purpose of the experiment and the intent behind the use of taboo words, and some conception of what the consequences for them will be

In the pilot study each taboo word was paired with a neutral word appearing in the same order, it was thereby possible to test differences between taboo and neutral words at successive points in the administration of the task The shock hypothesis would assert that taboo-neutral-word discrepancy scores should decline as the task progresses, so that taboo words appearing at the end of the task should be recognized nearly as readily as the paired neutrals To test for the presence of a decreasing trend a simple analysis of variance was employed Table 6 presents the differences between the pairs of taboo and neutral words at each of the eight positions The analysis of variance yielded a between-positions F of 12 46 significant beyond the 01 level We can conclude that the mean discrepancy scores for the eight pairs of taboo and neutral words are indeed dissimilar (i e , do not fall along a curve of zero slope), since they observably follow a decremental function, suggestive support for the shock hypothesis is afforded

The results of this experiment clearly support the hypothesis that the delayed reporting of taboo words is influenced by the interaction of the need for approval and task categorization The high-need-for-approval subject whose task categorization emphasizes social disapproval attempts to avert disapproval by prolonged avoidance, in contrast to the achievement orientation and more rapid recognition of the approval-motivated individual influenced by perceptual-speed cues Low-need-for-approval subjects did not display any effects of task categorization As a result of their lesser involvement and concern about the consequences of personal evaluation these subjects seem to have experienced less conflict

Our findings for low-need-for-approval subjects are perhaps clearer and require a less complicated explanation The taboo-neutral-word discrepancies displayed by these subjects are prob-

TABLE 6 Mean Difference in Taboo-Neutral-Word Recognition Thresholds at Each Position ($N = 28$)

Taboo-Neutral-Word Pairs	Position	Mean Difference
ounce *whore*	1	10 93
alert *penis*	2	5 50
prize *bitch*	3	1 86
yeast *screw*	4	5 89
Kotex table	5	4 11
raped cable	6	0 32
urine bread	7	1 07
breast magic	8	2 32

ably chiefly attributable to their initial surprise and "shock" at seeing "dirty words," a reaction that gradually gives way to relatively facile acknowledgement For approval-dependent individuals, as we have seen, the interpretation of the experiment greatly influences behavior As we speculated earlier, differences in the categorization of this experiment may have their origin in different childhood experiences—specifically, in parental punishment or acceptance of sexual curiosity and exploration Sears, Maccoby, and Levin (1957) portray this aspect of "the child's problem" as follows "When he learns to talk, he must master the control of sexual language He must not use

'naughty' words, nor ask certain questions, nor talk to other children nor outsiders about some of the things he learns in the bosom of the family—such things as the differences between boys and girls" (p 183) They point to the mislabeling, withholding of information, and general nonpermissiveness in the sexual area that characterize the child-rearing practices of many parents Perhaps, then, our "social disapproval" group was made up largely of subjects exposed to mislabeling, misinformation, and punitiveness in early sexual training Those high in approval dependence tend to perpetuate the effects of such training, while subjects low in the need for approval "learn differently" or rebel against these childhood prohibitions

It is interesting to note that for approval-motivated "perceptual speed" classifiers certain of the objective cues in the situation (the taboo words) were less salient as a result of their imposition of a "perceptual" and achievement meaning onto the task *None* of these subjects admitted withholding responses, in the postexperimental interview *80 per cent* of them did not mention the presentation of taboo words even after specific questioning, and they employed the least number of structurally similar alternatives and erasures Two possibilities suggest themselves First, we may speculate that these subjects experienced less punitive early sexual training, so that obscenity, while still socially disapproved, arouses less anxiety Second, these may be more achievement-oriented individuals Both these factors could, of course, operate in combination The failure of these subjects to mention the taboo words in the postexperimental interview is explicable if we recognize the difference between recording on an impersonal sheet of paper what one has seen and telling someone else face-to-face of this fact *Some* embarrassment may well have been felt by these girls

Along with the implications for the approval motive the results of this experiment also clarify the nature of perceptual defense as it is operationally defined and provide some suggestive answers to the long-standing question of the influence of response withholding A major component of variance in the differential recognition thresholds of taboo and neutral words seems to be the conscious inhibition of taboo-word responses

rather than perceptual blocking below the level of awareness Two major points support this assertion

1 Logically, from the theory of unconscious perceptual defense, those individuals showing high taboo-neutral-word discrepancy scores should achieve them by a process of unconscious blocking and distortion To the contrary, the highest discrepancy scores were attained by subjects who clearly consciously withheld many taboo-word responses Frank admission of response withholding was also observed in a large number of subjects

2 A perceptual-defense theory would also dictate the occurrence of blocking or distortion with equal intensity for each threatening word presented In other words, the concept of unconscious defense is usually taken to imply failure to discover changed conditions in the environment (from threat to absence or minimization of threat) Our pilot-study data and previous research (Bitterman & Kniffin, 1953, Lacey, Lewinger, & Adamson, 1953) indicate that recognition thresholds for subsequent taboo words decrease rapidly after recognition of an initial taboo word

If unconscious perceptual defense is thus an implausible explanation for the residual taboo-neutral-word discrepancies the question of how to account for the small remaining difference in recognition thresholds remains The explanation may lie in the "shock" hypothesis Introductory psychology students probably approach participation in psychological experiments with preconceptions which are almost certainly not completely allayed by the instructions given to them One preconception seems to be that psychologists are *really* interested in discovering something about the subject as an individual, i e, how abnormal or disturbed he is Given such a set it becomes the natural task of the subject to discover what the experimenter is actually after and to attempt to mitigate any unfavorable impressions In the perceptual-defense situation it would not be unexpected that many subjects would delay reporting taboo words until they had arrived at some conception of the experimenter's intent and had assured themselves that it would be all right (or at least inescapable) to acknowledge having seen "dirty" words

Those subjects who did this tended to be high in approval dependence

In demonstrating the joint effect of motive strength and expectancy (and, incidentally, their independence of each other in this experimental context) on perceptual defense, this experiment suggests the importance of assessing both variables in other experimental problems One obvious area of extension of the present findings is psychological testing Not only may a subject's goals influence his test behavior, as we have seen in research on the need for approval as well as other response-set conceptions (Couch & Keniston, 1960), but his expectancies regarding the consequences of his behavior and the interest and intent of the experimenter seem likely to be of major importance Thus, this study provides evidence highly consistent with the analysis of the testing situation that we presented in Chapter 2

part III Defensiveness and vulnerable self-esteem

8

Improvisation, public commitment, and attitude change

The compliance, conformity, and influencibility of approval-dependent persons ought to extend to the realm of attitudes and attitude change. We would clearly expect dependence on the approving support of others and fear of loss of their esteem to be associated with susceptibility to persuasion. Salman (1962) designed an experiment to test this proposition.

The most extensive and systematic exploration of personality factors in persuasibility in the literature to date is the work of the Yale Communications Group reported in *Personality and Persuasibility* (Hovland & Janis, 1959). One of the basic findings reported in this volume is that highly persuasible individuals appear to be characterized by low self-esteem, as manifested in feelings of social inadequacy, inhibition of aggression, and depressive affect. Replicating some previous findings of Janis (1954), Cohen (1959) demonstrated that subjects with low self-esteem were less likely to exert social influence and were more affected by the attempts of others to influence them. In another study reported in *Personality and Persuasibility,* Linton and Graham (1959) investigated the personality characteristics of three groups of subjects: opinion changers, nonchangers, and negative changers. The changers were found to be more visually field dependent, tended to be "other-directed"

on a questionnaire assessing this social and value orientation, held submissive attitudes toward authority, were anti-intraceptive on the California *F* scale (Adorno, et al , 1950), and gave Rorschach protocols characterized by weakness, passivity, immaturity, and repression In these studies, persuasibility emerges as a significant correlate of personality Moreover, the findings are cohesive those most easily persuaded to change their opinions seem to be characterized by a consistent set of attitudes toward themselves and toward themselves in relation to others The Cohen and Linton and Graham results can, thus, be linked by the concept of self-esteem

Our study of approval dependence has revealed a constellation of attributes in many respects not unlike the characteristics of the low-self-esteem subjects in the *Personality and Persuasibility* experiments There are, however, some inconsistencies and one glaring paradox—the difference in self-evaluative behavior between approval-seeking and low-self-esteem individuals We shall return later in this chapter to the question of self-esteem and try to unravel the threads of a complex theoretical problem But first there is the issue of how to induce attitude change—and, of course, the experiment itself

One approach to attitude change is to expose subjects to persuasive communications from a prestigeful source This was the basic procedure followed in many of the studies reported in the Hovland and Janis book Attitude change can also be induced by having each subject develop his own persuasive argument, this is the experimental design and method of inducing change of the improvised-role-playing studies of Janis and King (1954) and King & Janis (1956) In these experiments, each subject was required to rehearse, reformulate, and spontaneously present a persuasive communication in his own words Janis and King concluded that such improvisation is a critical factor in mediating changes in attitudes and is far more effective than the more passive receipt of the same content In effect, the subject "talks himself into" a new position

The improvisation and public delivery of a persuasive argument might be seen to have a rather different effect By improvising and presenting a set of ideas discrepant with one's

private beliefs, a state of inconsistency or conflict would be aroused. Reduction of the inconsistency or resolution of the conflict would follow from the subject's changing his attitudes more nearly to agree with his public persuasive behavior, or by some alternate means such as discrediting the situation or rationalizing his public commitment ("I was forced to do it"). This conceptualization of improvisation and the processes involved in producing changes in attitudes is similar to Festinger's (1957) theory of cognitive dissonance and to other related formulations (Heider, 1958, Newcomb, 1959; Osgood & Tannenbaum, 1955, Rosenberg, 1960), all of which are based on a central conception—the role of inconsistency or imbalance in the structure of attitudes and their related cognitive elements and the role of forces tending in the direction of consistency or balance. Thus, a public commitment to a position that one ordinarily would not consider produces a state of inconsistency or dissonance which may lead to private attitude change in the direction of the position one has been induced to take.

Salman predicted that persons dependent on approval who engage in role playing in the presentation of a persuasive appeal will demonstrate greater change in their attitudes. This hypothesis is based on the assumption that these individuals will experience greater inconsistency or conflict over their public avowal of a belief they do not privately hold. Because of their desire to be favorably evaluated, approval-motivated subjects are likely to be more constrained in an experimental situation and to be less likely to rationalize their behavior or to dismiss the experimental procedure as personally irrelevant. Thus, alternate means of resolving the discrepancy between a publicly avowed position and their private beliefs are less accessible to them. They are, then, compelled to be consistent with the public image they have created.

To assess the roles of personal commitment and improvisation in attitude change, the experimental design provided for three degrees of personal involvement or commitment to the issue. These were accomplished by the following role assignments: (1) "communicators," or active participants, who were required to improvise and to deliver an impromptu argument, (2) "receiv-

ers," who were passive recipients of the communication and whose assigned task was to evaluate the communicator's "leadership ability", and (3) "observers," who performed the same evaluative role but who, in addition, were physically separated from the other subjects by a one-way-vision mirror

Salman hypothesized that the greatest attitude change would occur among high-need-for-approval subjects given the role of communicator Less change was predicted for approval-motivated receivers and little or no change among high-need-for-approval observers Significant alteration of attitudes was not expected among low-need-for-approval subjects These predictions are graphically summarized in Figure 1

Early in the academic quarter, the *M-C* scale and a measure of attitude toward the desirability of revealing one's private feelings and personal problems to others were given to a large number of introductory psychology students. Male volunteers were later solicited for an experiment on "Leadership behavior," which was presented in the guise of a U S Air Force project

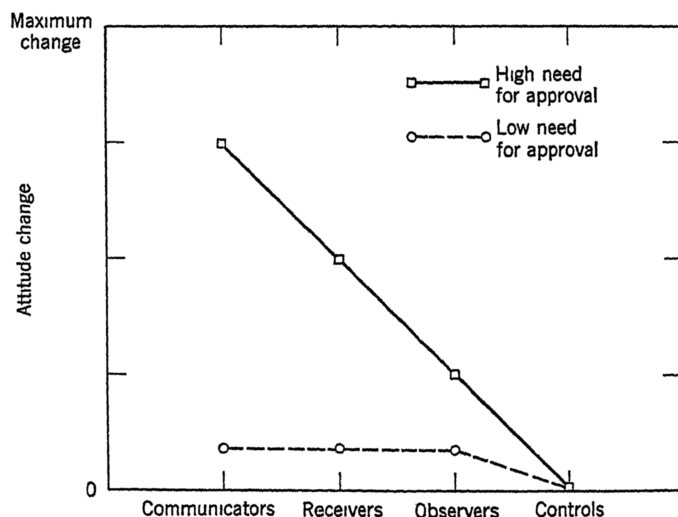


FIG 1 *Predicted attitude change for high- and low-need-for-approval groups in the various role assignments*

Of these volunteers, 150 participated in the experiment The design called for 3 subjects for each experimental appointment 1 to serve as a communicator, 1 as a receiver, and a third to take the role of the observer Due to scheduling difficulties and a shortage of males, however, a number of the groups consisted of only 2 subjects, the observer was omitted from these groups There were 57 groups that were actually scheduled and completed the experiment 23 groups of 2 subjects and 34 groups of 3 subjects It may be noted at this point that no difference in attitude change was found between the 2-subject and 3-subject groups

A test-retest control group was also included in the experiment 46 subjects merely repeated the attitude scale at an interval comparable to the period between pre- and postexperiment tests for the experimental subjects

The desirability or undesirability of revealing one's private apprehensions to others was deliberately chosen as the issue because of its personal salience Such an issue would be likely to evoke strong feelings one way or the other in most individuals Public commitment on an issue of personal importance should entail a greater degree of inconsistency and promote more pervasive efforts to reduce it In selecting a personally relevant and affectively linked attitude, Salman took a position contrary to that espoused by the Yale group (Hovland, Janis, & Kelley, 1953) In the Yale studies of attitude change, peripheral issues were presented to subjects in an attempt to avoid arousing motives to distort, suppress, or defend It is possible, however, that attitudes of marginal significance to the individual may undergo change simply because he has little investment in his position Significant attitude change on an issue of greater personal meaning is perhaps less likely to be based upon simple, minimally motivated compliance

The scale assessing attitudes toward personal revelation bore the innocuous title of "College Attitude Survey" and consisted of 8 relevant items embedded among 24 buffers Of the 8 items, 4 were worded in the "open" direction, advocating personal revelation, and the remaining 4 expressed the "closed" position,

TABLE 1 Scored Items on the Experimental Attitude Scale

1 It is common to find that those individuals who burden others with their inner feelings and problems are generally weak and soft-headed.
2 People who keep things to themselves are generally cold and aloof
3. Friendship must involve intimacy (the sharing of inner thoughts and feelings)
4 Nowadays, more and more people are prying into matters that should remain personal and private
5 Those who lay open their thoughts to others are just asking for ridicule and rejection
6 After having thought about one's problems, it is much better for a person to discuss his feelings and concerns with other people, rather than trying to reach a solution by himself.
7 The way to genuine friendship is to allow one's emotional life (even one's conflicts) to be known by another
8 Most people who talk about themselves do so only to gain sympathy.

against personal disclosure This was done to obviate the influence of an acquiescent response set Likert scales, ranging from strong agreement to strong disagreement, provided the response options on each item Pretesting indicated that each of the scored items yielded an approximately normal distribution of responses The 8 relevant items of the attitude scale are shown in Table 1

Each group to appear for the "Government Leadership Study" was met by the experimenter, an Air Force officer in uniform From each group, he randomly selected one of the subjects to accompany him to another room while the others remained in the waiting room This subject was given the communicator's role, and his specific assignment was presented as follows The officer was serious, intense, and emphatic as he outlined the task, stressing its life-or-death importance

This experiment is part of a large-scale government research project on *leadership* Our staff is attempting to identify and analyze the behavior and personality characteristics of people who we find have leadership capabilities We are approaching this study of leadership by way of the communication process Thus, we will be examining *speech* and also the methods people use in *delivering* verbal materials to others One of the aims of this study is to identify those elements of speech and the methods of delivery found in the verbal communications of effective *leaders* As you may know, an important part of leadership is a person's ability to influence another person's behavior or attitudes about some relevant issue You will be attempting to do this in this experiment In the past we have found that a major hindrance to effective leadership is the unwillingness of some men under a leader to disclose what is on their minds (like their feelings, problems, difficulties, etc) to others Instead, these people keep this kind of information to themselves—they are guarded and secretive about their problems and feelings To illustrate how important we find this issue to be, during World War II and the Korean conflict we found that many strategic bombing missions were failing because some crew member was disturbed about family troubles or other personal difficulties and therefore could not concentrate on the mission More important, the remainder of the crew was unaware of his problems and assumed that he knew what he was doing Since we find this to be an important issue we have chosen this as the topic for the persuasive argument in this experiment

Specifically, in this experiment you are going to take the role of a *leader* and will construct a short persuasive argument to be presented to another subject By way of your verbal presentation you will attempt to *influence* the other subject about the *importance and benefits of free expression of emotional feelings* The subject to whom you will present your communication will have received prior instructions He will be told to pay close attention to whatever you say But, he will remain as "objective" as possible and will not say or do anything himself (he will only *listen*) This is necessary because we wish to examine the speech and methods of delivery in "leaders" and a two-way conversation would bias these results This means that you will not know how you are doing in influencing the subject's attitudes about the issue Also, since you will not know whether the subject is pro or con with respect to the issue, it's best to assume that he *opposes* it and act accordingly After the experiment the subject will state how his feelings toward the issue have changed, *and* he will make an evaluation of your leadership potential at this time

The entire communication will last for about five minutes I will let you know when there are two minutes left so that you may start closing your argument or summarize some points at this time

It is suggested that you begin your presentation with some statement introducing the nature of the issue to the subject Then go on to discuss

the advantages and disadvantages as you see them Your task will be a lot easier and the results more effective if you are sure to include *your own personal experience* (drawing on events in your own life) in your presentation Remember, true personal experiences discussed with feeling and enthusiasm make the best persuasive argument

The communicator, thus, was required to improvise a persuasive appeal and to present it in a deliberate attempt to influence another While the communicator prepared his argument, the experimenter met with the other two subjects, giving them essentially the same instructions except for the material referring to a leader's ability to influence the behavior and attitudes of others These subjects were then assigned their roles One of them, the receiver, was told that he would be the face-to-face recipient of "a five-minute persuasive argument about the importance and benefits of free expression of emotional feelings" at the conclusion of which he would be required to evaluate the communicator's leadership potential on the basis of the efficacy of his verbal appeal The observer was given the same assignment but was required to carry it out in another room separated from the others by a one-way-vision mirror

The subjects were then escorted to their respective rooms and, as the communicator began his argument, the experimenter started a tape recorder, clicked a stop watch, and seated himself

When the communicator had finished his argument, the experimenter stopped the tape recorder and signaled to the subjects that the experiment was concluded Then, as if suddenly recalling something he had forgotten, he added

I have a reaction questionnaire to get at some of your feelings and reactions during the experiment I have to go get that But while I do, I'm also doing some work for the Psychology Department, they're constructing some new tests and they asked me if the subjects in the experiment had some extra time whether they could fill out some of these forms While I go get my reaction questionnaire, why don't you fill out these forms and I'll be right back (Pause) Some subjects have told me that they have taken one or the other of these tests before If you have, don't worry about it, just go on and fill them out again I'll be right back

The "Psychology Department's new test" was, of course, the experimental attitude scale, which was now given to the sub-

jects The experimenter gave them ample time to complete it before returning with the reaction questionnaires This measure was designed to assess the credibility of the experiment, the subjects' understanding of their assigned roles, their attitudes toward their own performance in carrying out the assignments given to them, and their feelings about the issue and any changes in their attitudes toward it

Responses to the reaction questionnaire indicated that the subjects were considerably involved in the experiment and highly concerned about their own performance (particularly the communicators) They accepted the experiment as a leadership study and did not connect the pre- and postexperimental attitude measures with the study of leadership

RESULTS

To form the high- and low-need-for-approval groups for the analysis of the results, scores on the *M-C* scale were dichotomized at the overall mean Change in attitude toward self-revelation was expressed by difference scores between the pre- and postcommunication testings The means and standard deviations of the experimental and control groups on the attitude scale are presented in Table 2[1]

In Table 3 the difference-score means and standard deviations of high- and low-need-for-approval subjects in each of the three experimental positions and in the control group are presented To determine whether the experimental manipulation was effective in altering subjects' attitudes and whether, as predicted, it was differentially effective, the mean amount of change of high- and low-need-for-approval subjects in each of the experimental positions was tested against the null hypothesis of a zero differ-

---

[1] In the experimental group high-need-for-approval subjects had a mean pretest score on the attitude scale of 20 34, in contrast to a mean of 23 25 for the low-need-for-approval subjects These pretest means differ significantly—beyond the 01 level Prior to the experimental manipulation, then, high-need-for-approval subjects were significantly more opposed to the open position advocated in the experiment than low-need-for-approval subjects

TABLE 2 Attitude-Scale Means and Standard Deviations for the Experimental and Control Groups

Group		Pretest	Posttest	Change Scores
Experimental	*N*	111	145	111
	M	21 73 *	24 14	22 34
	SD	5 64	6 21	4 98
Control	*N*	46	47	46
	M	21 70	22 04	20 26
	SD	6 16	5 09	4.63

* A constant of 20 was added to each score to avoid negative numbers

ence Examining first the high-need-for-approval communicator group, the obtained t ratio is 3 00, significant beyond the 001 level Very clearly, high-need-for-approval subjects who engaged in the preparation and public delivery of a persuasive appeal altered their attitudes in the direction they were required to advocate In the remaining analyses only among

TABLE 3 Change-Score Means and Standard Deviations

Group		Communicators	Receivers	Observers	Controls
High need for approval	*N*	22	18	18	18
	M	24 86 *	21 89	22.72 †	19 28
	SD	4 56	5 22	4 78	3 82
Low need for approval	*N*	21	20	12	27
	M	21 33	20 70	22 33	20 96
	SD	5 17	4 49	5 25	5 13

* Significantly different from zero change at the 001 level (two-tailed test).
† $p < 03$

high-need-for-approval observers was attitude change significantly greater than chance ($t = 2\ 42$, $p < 05$) Analyzing the data in this way, high-need-for-approval receivers and low-need-for-approval subjects in all three positions failed to demonstrate evidence of attitude change

A comparison of high- and low-need-for-approval communicators with their respective controls is shown in Table 4 Approval-motivated communicators changed in the advocated direction to a significantly greater degree than did their test-retest controls, while communicators less dependent on approval did not differ from their counterparts in the control group An analysis of the attitude-change scores of high- versus low-need-for-approval subjects disregarding experimental position revealed greater change on the part of the approval-motivated group ($t = 2\ 05$, $p < 05$)

The differential effects of the improvised persuasive communication may be examined in yet another way Table 5 shows the percentage of subjects changing in the advocated direction in each of the experimental groups Here, Salman analyzed his data to determine the *number* of subjects changing one or more points in the advocated direction By χ^2 tests the percentages of high-need-for-approval subjects in both the communicator and receiver positions who changed were significantly beyond chance expectancy

TABLE 4 MEAN DIFFERENCES BETWEEN THE CHANGE SCORES OF HIGH- AND LOW-NEED-FOR-APPROVAL COMMUNICATORS AND THEIR RESPECTIVE CONTROLS

	Communicator			Control				
Group	*N*	*M*	*SD*	*N*	*M*	*SD*	Diff	*t*
High need for approval	22	24 86	4 56	18	19 28	3 82	5.58	4 21*
Low need for approval	21	21 33	5 17	27	20.96	5 13	0 37	0 25

* $p < 001$, two-tailed test

TABLE 5 Percentage of Subjects Changing in the Advocated Direction

Group	Communicators	Receivers	Observers	Controls
High need for approval	86 *	78 †	56	28
Low need for approval	62	65	76 ‡	48

* $\chi^2 = 11\ 73$, $p < 001$
† $\chi^2 = 5\ 56$, $p < 02$
‡ $\chi^2 = 3\ 00$, $p < 10$

Salman then divided his subjects into those relatively opposed to the issue and those relatively in favor of it prior to the experimental manipulation This was accomplished by placing subjects in these two groups on the basis of the overall mean (21 72) on the pre-experimental attitude measure Subjects scoring above the pretest mean were considered to be relatively in favor of revealing their private feelings to others, while those with scores of 21 and below were regarded as being in opposition to personal revelation Table 6 shows the mean change scores of high- and low-need-for-approval subjects opposed to the issue in each of the experimental positions Analyzing the data of subjects initially opposed to the issue by testing against the null hypothesis of a zero difference, high-need-for-approval subjects in two of the three experimental positions demonstrated significant change and, at the 10 level, change in the third (observer) position The change scores of low-need-for-approval subjects in all positions were not significantly different from zero. A comparison of high- and low-need-for-approval communicators initially opposed to the issue yielded a t of 1 97, significant at the 07 level Disregarding experimental position, high- and low-need-for-approval subjects opposed to personal revelation differ significantly ($t = 2\ 23$, $p < 05$) in change

TABLE 6 Change-Score Results for Subjects Opposed to the Issue

Group		Communicators	Receivers	Observers	Controls
High need for	*N*	14	8	10	7
approval	*M*	26 36 *	24 25 †	23 90 ‡	20 86
	SD	3 32	4 26	6 08	5 24
Low need for	*N*	7	7	5	12
approval	*M*	21 86	22 57	22 40	22 58
	SD	5 58	2 89	4 33	4 93

* Significantly different from zero change at the 001 level.
† $p <$ 05
‡ $p <$ 10

scores Table 7 compares high-need-for-approval communicators in favor of the issue with those opposed to it Those approval-motivated communicators initially opposed to the issue demonstrated greater change than approval-dependent communicators initially more favorable to personal revelation Among low-need-for-approval communicators the same analysis yielded a nonsignificant *t* of 0 31 Among high-need-for-approval receivers those initially more opposed had greater change scores at the 10 level A comparison of high- and low-need-for-approval subjects in all experimental positions yielded a highly

TABLE 7 Mean Difference between Change Scores of High-Need-for-Approval Communicators Who Initially Opposed and Those Who Favored the Issue

Group		*N*	*M*	*SD*	Diff.	*t*	*p*
High-need-for-approval communicators	Pro	8	22 25	5 46	4 11	1 93	07
	Con	14	26 36	3 32			

significant difference in favor of approval-motivated subjects opposed to the issue and no difference between pro and con low-need-for-approval subjects These results are shown in Table 8

Before accepting the findings on the effects of initial position it is necessary to examine the possibility that these results may be due to ceiling or statistical regression effects The mean pretest *raw* score on the attitude scale of high-need-for-approval subjects in favor of the issue was 5 08, the comparable value for low-need-for-approval subjects was 5 88 This difference was nonsignificant The maximum possible raw score on the attitude scale was 24 (a maximum weight of 3 on each of 8 items). It is obvious that high- and low-need-for-approval subjects in favor of the issue were comparable on the initial attitude measurement and that on the average they had ample room in which to change their attitudes In point of fact, pro subjects changed little if at all, while the con group, especially those high in the need for approval, chiefly contributed to the observed change

The data of this experiment conform very closely to theoretical expectation Not only did approval-dependent subjects show greater attitude change, but they changed as a function of the degree of public commitment entailed by the role they played in the experiment and their initial opposition to, or acceptance

TABLE 8 Mean Differences between the Change Scores of High- and Low-Need-for-Approval Subjects Initially Favoring and Opposing the Issue

	Pro			Con				
Group	*N*	*M*	*SD*	*N*	*M*	*SD*	Diff	*t*
High need for approval	26	21 08	4 52	32	25 06	4 57	3 98	3 32*
Low need for approval	34	20.79	5 23	19	22 26	4 18	1 47	1 12

* $p < 01$, two-tailed test

of, the advocated position It is evident that more than compliance with the immediate demands of the experimental situation was involved If change in attitude was simply attributable to compliance or "prestige suggestion," little difference should have been observed between approval-motivated communicators and passive recipients Neither would the initial position findings have been expected on such a basis What actually happened, of course, was that high-need-for-approval passive recipients tended to show less attitude change than those approval-motivated subjects who engaged in improvisation and role playing

Playing the role of a sincere advocate of a position more extreme than their private beliefs seems to have created a state of inconsistency for subjects between their public persuasive behavior on the one hand and their own beliefs on the other. The discrepancy or inconsistency would be expected to be greater for subjects initially less favorable to the issue, resulting in attempts to reduce inconsistency Salman's findings clearly bear out this line of argument The fact that the greatest change was observed among approval-motivated communicators initially opposed to the issue suggests both a greater degree of inconsistency among these individuals and the existence of constraints limiting the possibilities for reducing it The situation for the subject—especially the approval-motivated communicator initially against the issue—was one of having openly committed himself in support of a position to which he was privately opposed As a consequence of his greater dependence on a favorable evaluation, the high-need-for-approval subject was less likely to be able to dismiss the persuasive task and the experimental situation as personally irrelevant or to devalue the advocated position and its credibility The major avenue remaining open to him to cope with a painful state of inconsistency was to change his privately held beliefs That inconsistency is indeed painful is attested to by the rigorous training of children in American culture to be consistent and not to say things they don't believe in And, of course, if one wishes to be well thought of, he had better observe such an injunction and be consistent—by changing his private views.

Less concerned with approval, subjects in the low group were, perhaps, more able to say to themselves in effect, "Well, it's just a psychological experiment, and I don't really care one way or the other" They seem, thus, to have experienced less demanding pressures to change Clearer specification of the means of inconsistency reduction among low-need-for-approval subjects—and, indeed, whether a state of inconsistency was aroused in this group at all—remains a problem for future investigation

High-need-for-approval receivers demonstrated an amount of attitude change roughly intermediate between communicators and observers The observers, physically separated from the interplay of the other two subjects, displayed minimal change in attitude It appears that the role of passive recipients involved a lesser obligation to concern themselves with the content of the persuasive argument, and these subjects were of course not posed with the conflict of publicly avowing a position discrepant with their private beliefs

An intriguing aspect of Salman's results is their apparent inverse comparability to the self-esteem findings reported in *Personality and Persuasibility* As we saw earlier, Cohen found that subjects who are most influenced by persuasive communications tend to be individuals lacking in self-esteem Salman demonstrated that subjects who are most persuasible are those who tend to describe themselves as being nearly perfect persons One might certainly ask the question, "Is the personal enhancement displayed by approval-dependent individuals a means by which they attempt to protect and to maintain a vulnerable self-image?" Perhaps, then, socially desirable responding and the dependence on the evaluation of important others which characterizes those persons whom we have labeled as approval-motivated represent a self-protective effort—an effort to avoid threats to self-esteem, anticipated social rejection, and the necessity of having to acknowledge that one believes oneself to be a bad and worthless person In his chapter in *Personality and Persuasibility,* Cohen developed a very similar analysis He characterized individuals who report very high self-esteem as using avoidant and self-protective defenses Among these persons, the defenses of repression, denial, and reaction formation

appear to constitute a basic means of contending with impulses that threaten to disrupt their relationships with significant other persons As Cohen suggested, these are defenses patterned on stereotypically acceptable behavior in American culture—defenses which, at the price of denial of certain of one's thoughts and feelings, might enable an individual to maintain viable relationships with others and to protect his shaky conception of himself Cohen argued that

> the use of avoidance defenses permits the individual to organize a cohesive and encapsulated self-picture After they have been developed as a means of handling inner impulses, avoidance defenses become behavioral modes themselves and determine the social reality to which the person exposes himself Consequently, persons with avoidance defenses can turn away from experiences which reflect unfavorably on their self-picture Persons whose defenses are more expressive may not be able to deal so selectively with external stimuli (p 117)

He went on to suggest a hypothesis which, in the light of Salman's findings, is of considerable significance to the concept of the approval motive

> A further implication of the differing reactions of highs and lows (in self-esteem) to experiences which potentially threaten the maintenance of their self-esteem concerns their susceptibility to influence in mass-communication situations Communications and/or communicators who threaten an individual's general picture of himself may be expected to influence persons of different levels of self-esteem differentially Such threatening appeals may be rejected more by those of high self-esteem than by those of low self-esteem On the other hand appeals which enhance an individual's self-picture might be accepted more by the highs than by the lows Thus, one determinant of acceptance of mass-communication appeals may be the differential responsiveness to threats of negative changes in the self-picture on the part of those with high and low self-esteem (p 119)

Leventhal and Perloe (1962) reported an experiment testing Cohen's self-esteem–fear hypothesis and found evidence to suggest that subjects characterized by high reported self-esteem tended to be more influenced by communications of an optimistic nature than by threatening persuasions, while a reverse pattern was found among subjects low in self-esteem

There is no evidence of threats to self-esteem in the persuasive content of this study However, if subjects believed they were

being evaluated on their ability as leaders, which clearly appears to have been the case, then considerable threat to the self-esteem of high-need-for-approval subjects would have been the outcome of failure to be effective persuaders In the attempt to avoid this painful eventuality, approval-motivated communicators would have had to take the persuasive role much more seriously, with a greater degree of personal involvement Thus, in their efforts to influence others, approval-dependent communicators would have been more compellingly put in the position of creating a discrepant or inconsistent state of affairs for themselves, resulting in alteration of their private attitudes Another possibility is that approval-motivated subjects achieved a vicarious enhancement of self-esteem by identifying with a position advocated by a prestige source, the U S Air Force We cannot definitely choose between these alternatives on the basis of the available evidence, but the considerable involvement of subjects in the "leadership" role suggests the validity of the former interpretation In any event, there appears to be an intriguing and suggestive relationship between the approval motive and the concept of self-esteem

These considerations are a prologue to the experiments reported in the next two chapters The studies to come are based on the suggestions advanced here and attempt to subject to experimental scrutiny the idea that the dynamics of approval dependence may involve the protection of a vulnerable self-image

9

Instigation to aggression, emotional arousal, and defensive emulation[1]

If approval-motivated persons are more dependent on the positive evaluations of others as a means of protecting a defensively enhanced picture of themselves, we might then expect that they would have greater problems in the recognition and expression of hostility. For approval-dependent individuals hostility would entail the threat of alienation from others, social rejection, unfavorable evaluations, and the resulting peril of damage to self-esteem. Furthermore, recognition and acceptance of the fact that one can hold hostile and resentful feelings toward others and the defensive belief that one is a nearly perfect person are sharply incongruent or dissonant cognitions in the pervasive value scheme of American middle-class culture. Accordingly, to maintain a defensive self-image would require denial of the nature of one's negative motivation.

A salient area of conflict and defense for the approval-motivated person would seem to lie in the awareness and expression of hostility. Along with the observations above this proposition is further supported by the fact that the *M-C* scale itself contains a number of items which pertain rather directly to the

[1] In somewhat altered form, the experiment reported in this chapter is published elsewhere (Conn & Crowne, 1964).

acknowledgement of hostility As an example consider the item, "I sometimes try to get even rather than forgive and forget"

Conn designed an experiment to test the hypothesis that high-need-for-approval individuals defend against the arousal of hostility by means of avoidant, repressive defenses in contrast to persons with a weaker approval need who are more able to express anger Following a line of reasoning elaborated and subjected to experimental study by Schachter and his associates (Schachter & Singer, 1962, Schachter & Wheeler, 1962), Conn assumed that anger is an emotional state involving, in addition to verbally and symbolically mediated cues, a state of physiological arousal as well Thus, for an individual to be able to define his emotional state as one of anger (or fear, depression, etc) both a set of cognitions signifying and appropriate to a state of anger and the physiological cues contingent upon arousal must be present Schachter and Singer (1962) develop this view of emotion as follows

> . an emotional state may be considered a function of a state of physiological arousal and of a cognition appropriate to this state of arousal The cognition, in a sense, exerts a steering function Cognitions arising from the immediate situation as interpreted by past experience provide the framework within which one understands and labels his feelings It is the cognition which determines whether the state of physiological arousal will be labeled as "anger," "joy," "fear," or whatever (p 380)

The Schachter studies attempted to assess the interaction of cognitive and physiological elements in emotional states by the artificial induction of physiological arousal (via the injection of epinephrine, a sympathetomimetic agent) and manipulation of the cognitions available to the individual to identify his emotional state The evidence from these experiments indicates that where appropriate cognitions are lacking, subjects tend to identify their emotional states in terms of the alternative explanations available to them and to demonstrate behavior consistent with the alternative cognitions

An issue seemingly raised by these studies concerns the problem of differential patterns of arousal in different emotional states and whether, if there are distinctive patterns, cognitive

elements in emotional states can be assigned a major influential role. In point of fact, research on the physiology of emotion has not clearly and unequivocally established consistent and unique patterns While certain experimental findings suggest the presence of differential patterning (Ax, 1953, Wolf & Wolff, 1947), there appears to be considerable overlap as between the emotional states of fear and anger (Buss, 1961)

For the present argument, however, the issue is not whether the discriminable patterning of physiological arousal in emotional states can be experimentally demonstrated but whether individuals can interpret their own bodily cues as a means by which to define and label an emotional state Schachter's results suggest that physiological cues by themselves do not provide a sufficient basis for the individual to interpret his emotional state

Pursuing this line of argument and extending it to the domain of ego defenses, Conn proposed that the effect of defenses is to block the cognitions defining a threatening emotional state He further argued that while defense against hostility (or any other emotional state associated with anxiety) affects verbal and symbolic cues, physiological arousal takes place nonetheless Thus, in a situation in which aggression is instigated, an initial recognition of cues leading to anger occurs but is subsequently defended against by repression, reaction formation, or other defensive processes In this assumption it is contended that defenses solve only part of the problem, namely, the individual's cognitive awareness of his anxiety-arousing motivational state Dollard and Miller (1950) developed a similar analysis in proposing that repression occurs when the verbal responses labeling a drive are inhibited [2] It follows, then, that given the imposition of a defense which fails to mediate the bodily cues associated with arousal, the individual will lack an explanation for a physiological state—why he feels "all stirred up "

[2] Dollard and Miller suggested two other cases of repression in which the drive is inhibited or weakened, but these are not relevant to the present discussion

This analysis would seem to be consistent with Freud's treatment of the concept of repression and repression-predicated defenses. The essentially cognitive effects of repression can be seen in the distinction which Freud drew between repressive defenses and emotional inhibition or ego restriction As Madison (1961) put it, emotional inhibition and ego restriction

> achieve ego-protection by preventing arousal of impulses rather than by distorting the conscious representation of already aroused impulses as in the repressive defenses
>
> In the defense of emotional inhibition ego-protection is achieved by preventing the dangerous inner impulse from developing into an active process When anger is inhibited, for instance, a person who has been unjustly treated will not feel angry No emotion develops despite provocation *This is different from becoming angry and distorting one's representation of one's anger* by reaction-formation, projection, or displacement in the respressive defenses (pp 38–39, italics ours)

Conn made one further assumption also made in the Schachter studies there is a need to explain one's feelings to oneself, and depending on the individual and the particular situational context some attempt will be made to resolve the inconsistency between a set of bodily cues and the absence of a satisfactory explanation for them A formal antecedent for this final assumption is Festinger's (1954) concept of a drive for self-evaluation by which to explain the tendency of persons to seek to resolve uncertainties and incongruities in their experience by social comparison

To summarize, Conn hypothesized that approval-dependent persons tend to use avoidant, repressive defenses against hostility These defenses act to block cognitive awareness of an emotional state and thus serve to prevent the arousal of anxiety. If, however, physiological arousal has taken place, high-need-for-approval individuals are posed with the problem of a set of physiological cues for which they lack defining cognitions For approval-motivated persons this state of affairs should lead, following exposure to a hostility-arousing situation, to the identification of their emotional state, and behavior consistent with that identification, in terms of the alternative cognitions available to them

EXPERIMENTAL METHOD

The experimental design basically involved two stages First was a hostility-arousing situation in which the subject was given clear provocation to anger by an experimental accomplice The second stage represented an attempt to supply the defensive subject with the opportunity to define his unverbalizable state in terms of an alternative set of cognitions—a different emotional state The alternative explanation selected was the emotional state of euphoria or hilarity—an emotional state strikingly different and sharply at variance with the immediately preceding instigation to anger The procedure, which was adapted from Schachter and Singer (1962), had the accomplice arousing the subject to anger behave in a highly euphoric manner, while the naive subject was rated for his degree of responsiveness to the accomplice's euphoria

Since the "base rate" of euphoric behavior of the high- and low-need-for-approval subjects is not known, and since the approval-motivated subject's euphoric responsiveness under these conditions might be explained simply on the basis of his greater influencibility, a control group was also required Subjects in the control group were paired with the accomplice prior to the euphoria situation in the performance of a neutral, non-anger-arousing task

At the conclusion of the euphoria situation, a set of postexperimental measures was administered to assess subjects' attitudes toward the accomplice and toward the procedures involved in the experiment

The 74 subjects who served in the experiment were all males, each of whom volunteered to serve with another subject, actually the accomplice Fictitious names were filled in for each available hour on the sign-up sheet to accomplish this necessary bit of deception. Three accomplices were used, all advanced undergraduate students in psychology, but credibly close in age and appearance to the subjects The subjects were alternately assigned to the experimental and control groups and to the three accomplices

The independent variable was, of course, the index of need for approval, which had been given to all introductory psychology students on the first day of classes There was, thus, no possibility that subjects might associate the scale and the procedures involved in the experiment

When the subject arrived in the office, he was asked by one of the secretaries to sit down and wait for the experimenter's appearance The accomplice consistently arrived a few minutes late, was similarly instructed, sat down, and began to engage the subject and the secretaries in conversation To maintain some consistency the accomplice played, from the very beginning, a kind of "hail-fellow-well-met" role Presently the experimenter arrived, informed the accomplice that another experimenter would shortly arrive to test him, and took the naive subject to another room

Several TAT cards were then administered [3] Following this, the subject, depending on the condition to which he was assigned, was taken either to the room employed for the hostility-arousal situation or to the one used for the neutral task He was informed that both subjects would participate in this phase of the experiment, and he was asked to wait while the experimenter went to determine whether the other "subject" was ready The experimenter shortly returned with the stooge who, upon entering the room, continued his sociable and locquacious role by commenting, "More stories?"

The subject and the accomplice were seated at opposite sides of a table divided in the middle by a large partition In front of each participant was a panel with two buttons, one red and one black A matrix indicating the outcomes of either black or red choices and a sign with the designation "Player 1" or "Player 2" were posted in front of each person The experimenter explained that the study was being conducted by the Air Force and that both players might win a considerable amount of money, depending upon how successful they were in playing a game The procedure, then, was a two-person, non-

[3] They were included to test some hypotheses regarding the relation of fantasy hostility to euphoria and are not relevant here

zero-sum game in which the payoffs were as follows If each player independently chose the black button, the payoff was $3 00 each The mutually competitive strategy, red-red, yielded a $ 10 payoff to each participant, and the outcome when one player chose red and the other pressed black was $5 00 to the red player and nothing to the black player These very high payoffs were intended to insure a high degree of involvement and to make the subject's failure to win as frustrating as possible

The experimenter explained the game and the payoff matrix, and a series of practice trials was conducted At this point, by prearrangement, a secretary interrupted to inform the experimenter that he had an urgent long-distance telephone call In annoyance and concern, the experimenter left the room, requesting that the subjects not discuss the game with one another and suggesting that they familiarize themselves with the payoff matrix After a moment, the stooge leaned around the edge of the partition and said to the subject, "Say, I've got it I've figured out how we can win the most money All we have to do is both press black each time and we'll both win $3 00 each time After all, it's no skin off his teeth—it's the Air Force's money, isn't it?" The accomplice argued for his proposition until the subject agreed, concluding with, "OK, I'm counting on you to play black like I will" Occasionally, the subject himself initiated a proposition to which the accomplice agreed only if it involved a cooperative strategy If not, the accomplice argued as above After a delay of approximately a minute and a half, the experimenter returned, made conspicuous a roll of bills and a box containing dimes, and explained to the subjects that the game would consist of five trials The game was then begun, and on each of the five trials the accomplice now pressed red, violating his secret agreement with the subject Following each trial, the players were paid off, with the frequent result that the subject played cooperatively and received nothing, whereas the stooge, capitalizing on his ruse, received $5 00 After each trial, the experimenter announced each participant's choice of strategy and ostentatiously made the payoff

In the control group, the subject and the accomplice were taken to a room, seated at a table, and given a set of pictures

on which they were independently to make forced-choice ratings on a series of bipolar adjectives This rating task, of course, bore no relationship to the study and was included solely for the purpose of providing the subjects with a period of time with the stooge comparable to that of the experimental group

At the conclusion of the game (or the neutral task for the control group) the subject and the accomplice were led down the hall to another room, where they were instructed to wait until the apparatus could be readied for the next part of the experiment During the trip down the hall, the experimenter walked considerably ahead to give the experimental-group subject an opportunity to express his reaction to the accomplice's perfidy If the subject said anything to the stooge about his violation of the agreement, the stooge laughed and said, "Well, that's the way it goes"

The small room in which the subject and accomplice were assigned to wait had a sign on the door saying "Child Therapy Room" and on a table in one corner was a diverse assortment of objects—toys, paper, rubber masks, children's drawings, a frisbee (plastic disk), crayons, etc Opposite the door was a one-way-vision mirror permitting observation of the entire room There were only two chairs in the room, and the accomplice was instructed to leave the subject the chair directly facing the one-way mirror

The accomplice gave the experimenter enough time to station himself on the other side of the mirror and then began his euphoric routine, declaring, "Boy, I feel great today, this is one of my better days" He next told some atrociously bad jokes Laughing uproariously, he slapped his leg and doubled over in mirth As the routine developed, the accomplice crumpled up wads of paper and shot at the wastebasket, he then elaborated an imaginary basketball game in which he attempted to get the subject involved He also made paper airplanes, tried to get the subject to do the same, encouraged the subject to tell him jokes, and threw the plastic disk across the room where it landed in the wastebasket with a resounding clang The ac-

complice again urged the subject to tell him a joke and concluded with one of his own, ending the routine

Throughout the entire sequence, the accomplice invariably maintained his good humor Negativism or hostility on the part of the subject was greeted by such comments as, "Come on, live a little!" or "What's the harm in a little fun?"

The stooge's routine was divided into 17 different steps or units, each of which called for a rating The sequence and content of the euphoria procedure were constant from subject to subject, and the only variations introduced were those occasioned by the initiation of activity by the subject Such occurrences were greeted enthusiastically by the accomplice Having responded warmly to these ventures when they occurred, the accomplice then attempted to return to the stage where he left off

The experimenter and a second rater observed and listened to the entire sequence from behind the one-way mirror, rating the subject's behavior for each of the 17 units Two ratings of euphoria were made one of facial activity and the other of actual physical participation with the accomplice The ratings were made on 6-point scales ranging, in the case of facial expression, from 1, a negative response (frown, look of disgust, anger, or displeasure, looks away from the accomplice or takes out something to read) to 6, a big laugh (belly laugh, or explosive sounds of greater volume accompanied by marked body movements, e g , doubling up, throwing up hands, slapping leg, holding mouth or side) The low point of the activity scale (1) included the following tries to get the accomplice to stop his routine either verbally ("Act your age" "Why don't you sit down?" "Who are you trying to impress?") or physically, by direct contact with the accomplice or by appropriating and retaining the accomplice's play objects A rating of 6 was reserved for activities in elaboration of the accomplice's (execution of a behind-the-back shot at the basket, playing with one of the articles on the table, starting a game of catch with the accomplice)

When the stooge had completed the routine the experimenter re-entered the waiting room and explained that the apparatus to be used in the next part of the experiment was broken and

could not be repaired in time The subject was taken to another room to complete a postexperimental questionnaire which contained some self-descriptive items and also a series of items on which the accomplice was to be rated At the same time the stooge filled out a scale describing the subject's reactions to him and reported verbatim any comments of the subject pertaining to the game Finally, a postexperimental interview was conducted with the subject to determine his reactions to the experiment and to the stooge Concluding this final stage of the experiment the experimenter informed the subject of the deceptions practiced on him—that the other subject was actually an accomplice and that the purpose of the experiment was to assess how people would react in complex situations such as the one to which he had just been exposed

RESULTS [4]

To test the major hypothesis the high- and low-need-for-approval groups were separately compared in the experimental and control conditions High-need-for-approval subjects in the experimental group were also compared to their counterparts in the control group, and a similar analysis was conducted for low-

---

[4] As a first step the postexperimental interviews were examined to identify those subjects who caught on to the experiment Eight subjects suspected that the other "subject" was actually an accomplice, and they were discarded Since three accomplices were used, it was necessary to determine whether they elicited different amounts of euphoria To this effect the mean total euphoria scores of the subjects run by each stooge were compared separately for the control and experimental conditions by means of t tests No differences were found

The interrater reliability of the ratings of euphoric behavior was determined on a pilot sample of seven subjects In a rating-by-rating comparison over the 17 units the two observers achieved 79 per cent agreement on the facial scale and 89 per cent agreement on the activity scale Disagreements of more than one scale point were extremely rare on both scales In the final data the ratings of the two observers were averaged for each subject, and scores on the facial and activity scales were combined to yield a total euphoria score

The high- and low-need-for-approval groups were formed by dichotomizing scores on the *M-C* scale at the overall mean

TABLE 1 Mean Amount of Euphoria of High- and Low-Need-for-Approval Groups in the Experimental and Control Conditions

	Need-for-Approval Groups						
	High			Low			
Condition	*N*	*M*	*SD*	*N*	*M*	*SD*	*t*
Experimental	16	3 78	37	18	3 19	67	3 24*
Control	16	3 45	49	16	3 58	60	0 65
t		2 12 †			1 77 ‡		

* $p <$ 01, † $p <$ 05, ‡ $p <$ .10, two-tailed tests

need-for-approval subjects. These results are presented in Table 1

To summarize these findings, high-need-for-approval subjects in the experimental condition were significantly more euphoric than those in the control group, and low-need-for-approval subjects in the experimental condition tended to be less euphoric than their corresponding controls Between-group comparisons in the control condition yielded no difference, as expected, while in the experimental condition high- and low-need-for-approval subjects differed at a highly significant level Figure 1 displays the results of these analyses

Conn then rescored the *M-C* scale, computing a new index based only on those items directly pertaining to the admission or denial of hostility With these new scores he repeated the analyses detailed above, with the result that no enhancement of prediction was obtained The major findings on the amount of euphoria thus appear to reflect more than the admission or denial of hostility on the scale itself and seem instead to be related to the more general disposition to give or not to give socially desirable responses

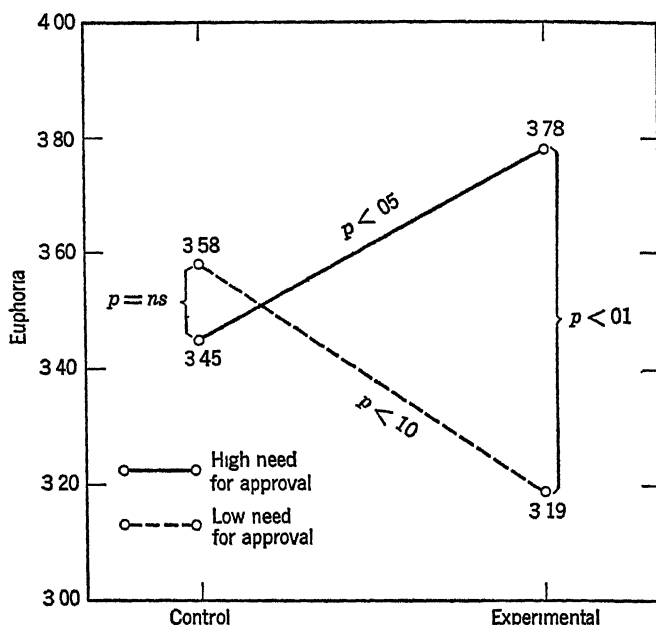


FIG 1 *Mean amount of euphoria of high- and low-need-for-approval groups in the experimental and control conditions*

As a check on the relationship of the facial and activity indices correlations were computed between these two scales separately for the control and experimental groups In the control group the correlation was 67, while for experimental subjects the correlation was 40 The lack of a perfect or near perfect correlation between these two scales is expected Smiling at someone else's antics does not necessarily mean that one will engage in similar activity with him Another way of putting it is that the activity scale should make an independent contribution above and beyond that of the facial index Nevertheless, the two measures are moderately correlated and appear to reflect different aspects of the same variable

The postexperimental questionnaire was divided into three major sections: a series of self-rating items, a nine-item forced-choice adjective scale on which the "other subject" was to be

rated, and, last, two open-ended questions calling for the subject to write out his evaluation of the accomplice. The self-rating items were mainly included to disguise the purpose of the scale.

Of the nine items in the forced-choice adjective scale, seven pairs clearly involved the choice of favorable or unfavorable ratings of the accomplice, i.e., friendly-unfriendly, dishonest-honest, hateful-likeable, and so forth. A total unfavorability rating of the accomplice was computed for each subject by summing over these seven items. These scores, then, reflect the willingness of the subject to characterize his coparticipant negatively and may be presumed to reflect the degree of the subject's anger toward the accomplice.

High- and low-need-for-approval subjects were then compared in the experimental and control groups, as in the test of euphoric behavior. Table 2 presents these findings. In the experimental group high- and low-need-for-approval subjects differ significantly in their characterizations of the stooge on the adjective scale, with the lows giving the expected negative evaluations. None of the other comparisons was significant, al-

TABLE 2. Mean Ratings of the Accomplice by High- and Low-Need-for-Approval Groups in the Experimental and Control Conditions

	Need-for-Approval Groups						
	High			Low			
Condition	*N**	*M*	*SD*	*N**	*M*	*SD*	*t*
Experimental	12	1.50 †	1.09	14	3.07	1.27	4.13‡
Control	16	2.00	1.32	16	2.31	1.49	0.50
t		1.10			1.51		

* Because of insufficient time 4 high- and 4 low-need-for-approval subjects did not complete the ratings of the accomplice.

† The higher the score the more negative the rating given to the accomplice.

‡ $p < .001$, two-tailed test.

though it is interesting to observe that the trend among high-need-for-approval subjects is in the direction of more favorable evaluations in the experimental condition, while the reverse trend was observed among the lows In the experimental group the highs, while objectively exposed to the same provocation to anger as the lows,[5] gave significantly less negative evaluations Such constraint was clearly not observed among low-need-for-approval subjects

The two open-ended questions which asked for an evaluation of the accomplice were "What kind of a person would you say he is?" and "Would you expect him to be successful socially and monetarily?" Responses ranged from such an extreme as "I think he's nuts Why, he was in there making paper airplanes What a screwball that guy is He doesn't act grown up" to "A very enjoyable person Great company An extrovert" These responses were given to three raters to be blindly scored 0, 1, or 2 for absence, mild expression, or strong expression of hostility Eighty-eight per cent agreement was obtained The values assigned by each of the three raters were pooled to yield an expressed-hostility index for each subject As in the preceding analysis of the adjective scale the only significant difference was that between high- and low-need-for-approval subjects in the experimental group, with the lows expressing more hostility than the highs ($p < 05$) On this measure the distribution of hostility scores of highs in the experimental group was quite similar to that of the control group, and in the control group high- and low-need-for-approval subjects did not differ Thus, the significant difference between highs and lows in the experimental group is clearly occasioned by the greater tendency of the lows to give critical—and, in some cases, openly hostile—evaluations

The results of this study confirm the predictions High- and low-need-for-approval subjects were clearly divergent in the amount of euphoric behavior in the experimental condition, and approval-motivated experimentals were much more eu-

[5] The mean number of black (cooperative) plays in the five-trial game for highs was 2 38, for lows, 2 72 ($p = \text{ns}$) Both groups were equally (and considerably) taken in by the stooge

phoric than their controls There was an opposite trend among low-need-for-approval subjects, with lows in the experimental condition tending to be less euphoric than lows in the control group Further, on the two postexperimental indices approval-motivated subjects in the experimental group characterized the stooge more favorably and expressed less hostility toward him than did low-need-for-approval experimentals

This experiment was predicated on a lengthy chain of assumptions about the nature of emotional states and repressive defenses—a chain, however, with some empirical support from the findings of the Schachter studies The extension of Schachter's analysis of the roles of cognitive and physiological elements in emotional states to the problem of ego defenses received clear support in the present data Given the following facts—the equivalent amount of euphoric behavior of high- and low-need-for-approval subjects in the control condition, the objectively equal provocation to anger of the two groups in the experimental condition, the different susceptibility to euphoria of highs and lows following hostility arousal, and the very low incidence of criticism and hostility in the postexperimental verbal reports of approval-dependent experimentals—it is difficult to invoke a plausible alternative interpretation

The complex nature of this experiment was necessary in order to get away from the obvious problems entailed by reliance on verbal report in the assessment of ego defenses By itself the fact that an individual does not verbally aggress following instigation to aggression cannot be taken as an indication of the presence of repression or repression-based defenses Clearly, a more parsimonious explanation in this case would be that the subject has simply decided to withhold a socially less acceptable response An experimental situation is very likely to involve constraints on the expression of aggression, particularly if the subject feels that he is being personally evaluated Thus, response withholding is no less likely here than in the perceptual defense situation we discussed in Chapter 7 However, alongside the greater reliance of approval-motivated subjects on the alternative cognitions supplied to them, their failure to react to the dirty trick played on them in their evaluations of the

experimental accomplice can be seen as providing supplementary evidence of the existence of repression.

The susceptibility of approval-motivated subjects to the stooge's euphoria is plausibly accounted for as the result of repression, and given the analysis of ego defenses presented here this is perhaps the simplest explanation. Another possibility, however, deserves exploration. Following the hostility-arousing situation the reaction of approval-dependent subjects to the stooge was responsive, friendly, and appreciative. Their behavior was quite the opposite of the kind of response they were provoked to make and that low-need-for-approval subjects did in fact make. What is suggested here is that high-need-for-approval subjects might have engaged in reaction formation, tending to respond affiliatively as a defense against their unacknowledgeable anger. Thus, in selecting the emotional state of euphoria as the alternative cognition to supply to subjects perhaps Conn simply hit upon the kind of behavior that approval-dependent individuals would naturally employ in defending against hostility. This experiment can provide no answer to the question of whether repression or reaction formation is the primary defensive process. Repeating the experiment using another emotional state such as depression or fear, or employing a measure of euphoria not dependent on the face-to-face interaction of the subject and an accomplice, would provide a means to resolve the issue. At this point, however, it seems safe to conclude that repression, or a repression-predicated defense, is associated with approval dependence.

From these results two modal resolutions to the problems of aggression and defense emerge in a suggestive formulation. The pattern displayed by high-need-for-approval persons involves the imposition of repression or reaction formation, defenses which serve to block awareness of the verbal and symbolic cues of anger. As a function of their anxiety and defensiveness these persons appear to have accounted for their aroused state (for which, as a result of defensive blocking, they had no adequate explanation) by being "carried away" and concluding with the accomplice, "Boy, I feel great today!" In dramatic contrast is the behavior of low-need-for-approval individuals, who tended

to respond with considerably more negativism to the individual who unconscionably violated their agreement Thus, the other resolution involves awareness of anger and its more frank expression

To recapitulate, Conn's findings provide striking evidence that the conforming, submissive, and influencible behavior of approval-dependent individuals is associated with defense against hostility

10

Further studies in defensiveness and vulnerable self-esteem

Persons dependent on the approval of others seem to maintain defensive, encapsulated pictures of themselves. Another piece of their defensive armor appears to be the use of repressive defenses against hostility. This is a vast extension of the meaning of the concept of need for approval, and it suggests a host of possibilities to examine. The experiments to follow test some of the issues raised by the findings on defensiveness and vulnerable self-esteem.

A possibility immediately suggests itself if one asks the question, "To what kinds of threat would approval-dependent persons be especially susceptible?" If the self-esteem-and-defensiveness hypothesis is correct, psychotherapy would tend to create for approval-motivated individuals a painful conflict between the strong need to maintain and defend their vulnerable conceptions of themselves and the necessity of discussing personal feelings, impulses, and problems in a meaningful way as a means of altering disturbed patterns of behavior. Psychotherapy, thus, would pose a severe threat to the self-esteem of approval-motivated persons. In the early hours of therapy, before an enduring and trusting relationship with the therapist has been established and before the patient can really believe that the pain of self-disclosure will indeed result in positive changes

in his life and his feelings about himself, the need to defend his shaky self-conception should tend to result in strong resistance Certainly one form of resistance (as a matter of fact, a manifestation of resistance which Freud observed and commented on as early as 1913) is abruptly to terminate psychotherapy Faced with the prospect of self-revelation and the threat this poses to his defensive self-image, the patient convinces himself that psychotherapy is really not worthwhile or that in a very few hours he has actually so improved that he has accomplished what he came for Following the logic of the defensiveness-and-vulnerable-self-esteem hypothesis, we arrive at a rather paradoxical prediction. despite their greater dependence on a favorable evaluation, patients high in the need for approval should find the demands of psychotherapy for personal revelation threatening, and this should lead to an avoidant means of resistance—early termination.

In the test of this hypothesis the *M-C* scale was given to a large number of patients at an outpatient psychiatric clinic [1] Throughout the interval spanned by this research almost every patient seeking help at the clinic was admitted to psychotherapy, which with only rare exceptions consisted of weekly therapeutic hours. A majority of the patients were given neurotic or character-disorder diagnoses, although many were classified as psychotic or prepsychotic

The patient sample was composed of two groups In the first group were 23 patients, 13 males and 10 females, who terminated psychotherapy during the months of October and November 1960 The second group—62 patients, 26 males and 36 females—was a replication sample terminating sometime during May, June, or July 1961

In addition to completing the *M-C* scale each patient rated his own improvement on a 9-point scale from improved (1) to unimproved (9) The therapists were required to rate each of their patients on a slightly modified version of the Seeman (1954) Case Rating scale Of interest in this study were five of

[1] This study is from Strickland and Crowne (1963)

the Case Rating scale items #5, the therapist's estimate of the patient's attitude toward him during psychotherapy, #6, the therapist's attitude toward his patient, #7, the degree of personal integration versus defensiveness of the patient, #9, the therapist's estimation of the patient's satisfaction with the outcome of therapy, and #10, a rating of therapeutic outcome In contrast to the original, which used 9-point scales, the range of values for this version was from 1 to 16 Other variables included education and occupational level, which were combined in a fashion similar to the procedure followed in the Hollingshead Index of Social Position (Hollingshead & Redlich, 1958) The number of hours of therapy at termination was the major dependent variable

The *M-C* scale was mailed to both groups of patients within a three-month interval from the actual date of termination The rate of return on the scale was approximately 70 per cent, and a comparison of the means and variances on the *M-C* scale of the two groups of patients with the mean and variance of another group given the scale at the clinic revealed no differences The distribution of hours of therapy of the two groups covered a very wide range and included a large number of early terminators who would be most likely to fail to comply with further requests from the clinic From this evidence it appears unlikely that a major degree of bias affecting either the independent or dependent variables was introduced into this study by administering the *M-C* scale after termination

The self-rating scale of own improvement in therapy was mailed to the patients along with the *M-C* scale Of the 85 patients, 63 completed this measure

The therapists completed the Case Rating scale at varying times, depending on the date each patient began treatment, between November 1960 and May 1961 The therapists' ratings were uniformly made while the patients were in treatment At the time the ratings were made the therapists could have had no reliable knowledge of the patients' ultimate dates of termination These ratings, then, could not have been contaminated by

TABLE 1 Distribution of the Number of Hours of Psychotherapy for High- and Low-Need-for-Approval Patients

Hours of Therapy	High Need for Approval (N = 43)	Low Need for Approval (N = 42)
71 + (to 284)	5	11
61–70	1	1
51–60	1	5
41–50	0	5
31–40	7	2
21–30	10	9
11–20	11	5
1–10	8	4

the possibility of negative attitudes toward early-terminating patients[2]

Before the major analyses the data were examined for sex and age differences. None were found except for the sex difference in patients' ratings of their own improvement reported below, and the analyses were accordingly carried out without regard to sex and age

To accomplish the test of the early-termination hypothesis patients' scores on the *M-C* scale were dichotomized at the overall mean to yield the high- and low-need-for-approval groups The distribution of the number of hours of therapy of these groups is shown in Table 1 By a Mann-Whitney U test high- and low-need-for-approval patients differ significantly in hours of therapy ($z = 2\ 84$, $p < 005$[3]) From the distribution of therapy hours in Table 1 it is clear that highs are underrepre-

[2] None of the therapists had any knowledge of their patients' need-for-approval scores, in fact, few of them had even a general knowledge of the nature of the research

[3] Two-tailed test

sented above the median hours of therapy and overrepresented below the median hours of therapy Low-need-for-approval patients tend to be more equally split above and below the median At the extremes of the distribution, however, the lows are clearly overrepresented above and underrepresented below the median Differences of approximately the same magnitude were found in the two samples (the first and second groups), giving some evidence for the replicability of the major finding

For the 30 patients in the second group on whom the therapists completed the Seeman scale, correlations were computed between these ratings and the *M-C* scale, the number of hours of therapy, and the index of social class These correlations and the intercorrelations of the five Case Rating scale items are presented in Table 2 Approval-dependent patients tended to receive more negative ratings on each of the five scales The most striking of these correlations is that between the need for approval and therapists' ratings of personal integration versus defensiveness ($r = -67$) [4] The remainder of the correlations are consistent but fail to reach a conventional level of significance

Each of the Seeman scales correlated moderately and positively with the number of hours the patient remained in therapy The index of social class was found to be unrelated to the *M-C* scale and negligibly correlated with length of stay in therapy

Since the patients had been in therapy for varying lengths of time when the therapists' ratings were made,[5] it is conceivable that the approval-motivated group was rated lower simply as a function of their tendency to remain less long in psychotherapy Their negative ratings could simply reflect the fact that their therapists knew them less well To test this possibility partial correlations were computed between the *M-C* scale and the therapists' ratings with the number of hours the patient had been in therapy at the time of rating partialed out Table 3 shows

---

[4] On the defensiveness scale the lower the score the higher the defensiveness

[5] The pressures of other research made it impossible to obtain the therapists' ratings after a standard interval for all patients

TABLE 2 INTERCORRELATIONS OF HOURS OF THERAPY, NEED FOR APPROVAL, AND THERAPISTS' RATINGS

Variables	1	2	3	4	5	6	7
1 Need approval	68 *						
2 Hours of therapy	− 28 †						
3 Patient's liking/respect for therapist	− 18	23					
4 Therapist's liking/respect for patient	− 35	40 ‡	66 §				
5 Personal integration vs defensiveness of patient	− 67 §	57 §	27 ‡	38 §			
6 Patient's satisfaction with therapy	− 32	22	60 §	46 §	55 §		
7 Amount of improvement in therapy	− 34	.45 ‡	46 §	50 §	63 §	79 §	
		$N = 83$ (columns 2–6)					
8. Social class	− 07	.19	27	00	28	36 ‡	07
	$N = 30$ (columns 1–7)						

* Test-retest reliability, 5–6 month interval during which patients were in therapy ($N = 24$)
† $N = 85$, $p < .01$.
‡ $p < 05$.
§ $p < 01$

TABLE 3 PARTIAL CORRELATIONS BETWEEN NEED FOR APPROVAL AND THERAPISTS' RATINGS

Therapists' Ratings	Partial Correlation with Need for Approval
Patient's liking/respect for therapist	– 11
Therapist's liking/respect for patient	– 29
Personal integration versus defensiveness of patient	– 63 *
Patient's satisfaction with therapy	– 25
Amount of improvement in therapy	– 24

Note $N = 30$
$^{*}p < 01$

these correlations The relationships between the need for approval and the therapists' ratings maintain approximately the same magnitude irrespective of the amount of time the patient had been in therapy at the time he was rated

There were no differences in the diagnoses of high- and low-need-for-approval patients Psychiatric diagnosis was further unrelated to the other variables of the study

The terminal ratings of own improvement were correlated separately for males and females with the *M-C* scale For males there was no relationship between ratings of own improvement and the need for approval ($r = 01$) A significant correlation was obtained for females ($r = -39$), with those more approval-dependent tending to rate themselves as more improved The own-improvement ratings correlated – 16 for males and – 26 for females with hours of therapy

A question we have not considered is the effect, if any, of psychotherapy on the need for approval Would scores on the scale change after a period of time in therapy? To provide an answer to this question the *M-C* scale was readministered to a

sample of 24 patients who had been tested five months earlier and remained in treatment The test-retest correlation was 68 There were no systematic shifts in the number of patients changing in score from the first to second testings nor in the magnitude of their changes in score Virtually as many shifted up as down and in about the same degree

Approval-dependent patients terminate psychotherapy much earlier than those less approval dependent To accept the hypothesis that the earlier termination of the high-need-for-approval group represents avoidance and defensiveness it is first necessary to exclude the possibility that they are just less disturbed and less in need of help than low-need-for-approval patients As established by therapists' ratings, approval-dependent patients appear to be more defensive Moreover, they tend to be judged by their therapists as less personally liked, less satisfied with the progress of therapy, and less improved in treatment Finally, no systematic differences were found between the high- and low-need-for-approval groups in type or severity of diagnosis It would appear that the approval-motivated group did not terminate earlier as a result of progress and improvement; their early termination thus takes on the character of resistance and defensiveness

These findings fit with the conceptualization of the approval-dependent person as one who is concerned with protecting a vulnerable self-image For the approval-dependent patient psychotherapy means a dilemma—to give up his defensive self-conception or defy the therapist The outcome of such an avoidance-avoidance conflict is leaving the field It is also possible that these patients fail to receive important affectional and dependency gratifications from their therapists—gratifications which might alter the balance of the conflict A corollary but not incompatible interpretation is that the approval-motivated individual is apprehensive about possible social criticism for seeking therapy and the implicit admission entailed "I'm crazy"

In this study both time pressures and another large-scale research project being conducted at the same time necessitated administering the *M-C* scale after the conclusion of psycho-

therapy This raises a crucial methodological issue whether patients' need-for-approval scores were affected by participation in psychotherapy If this were so, these findings would be more parsimoniously interpretable in terms of the influence of therapy on the questionnaire As detailed above, an attempt was made to check on the possible effect of therapy on the scale by doing a long-term test-retest analysis No systematic or even appreciable changes were found The nearly identical pre- and posttest means (10 92 versus 10 42) [6] suggest the strong resistance to change of the need for approval High on any list of procedures to alter approval dependence would certainly be psychotherapy But despite its pervasive emphasis on the disclosure of feelings and frankness in self-evaluation, those high on the scale tend to maintain their defensive images

In addition to their implications for the vulnerable-self-esteem-and-defensiveness hypothesis, these findings also have some suggestive value regarding the problem of a model for psychotherapy The verbal-conditioning paradigm has captured wide attention in the recent literature as an analogue of the process of psychotherapy (*cf* Krasner, 1958) According to the verbal-conditioning model psychotherapy is a process in which the therapist subtly reinforces changes in the patient's verbal behavior Success in therapy means that these changes generalize to verbal and behavioral changes in the patient's real life outside of therapy If we can assume that what is measured by the *M-C* scale is comparable in patients and college students, then the greater amenability to verbal reinforcement of approval-motivated individuals in a verbal-conditioning situation (Chapter 4) and their tendency to terminate therapy early and with less improvement pose a real problem for the model Whatever the other merits or demerits of the verbal-conditioning analogue, it certainly fails to account for individual differences and the meaning of the situation as it is perceived by the individual It is one thing for the approval-dependent person to be more influenced by the subtle reinforcers of an experimenter

[6] These means do not differ from the mean of the remainder of the patients in this study

It is evidently quite another thing to ask him to surrender his defensive conception of himself.

CONFORMITY AND CONFLICT

Let us turn to a slightly different problem now, one that involves some theoretical refinements in addition to implications for the defensiveness hypothesis We have relied so far simply on a concept of need in accounting for behavioral directionality The problem is a much more intricate one, however, when viewed from the perspective of Social Learning Theory In Social Learning Theory, as we saw from the brief review in Chapter 2, the probability of occurrence of a given behavior is considered to be a function of two variables the relative value of, or preference for, certain reinforcements, and the individual's subjective probability or expectancy that these reinforcements can be attained Suppose we were to consider the need-for-approval scale as an index of the relative value of approval and affectional reinforcements (To consider it in this way does not mean that we need to abandon the vulnerable-self-esteem hypothesis) If we now had a measure of generalized expectancy of success as a result of one's own abilities, we ought to be able to make more refined predictions

In Social Learning Theory, conflict occurs and maladjusted behavior is probable when a goal is positively valued but the individual has a low expectancy of goal attainment For individuals with a high need for approval but relatively low expectancy of success (that is, low freedom of movement), we should expect to observe avoidant, defensive, and self-protective behavior

We used a social conformity situation to test this formulation [7] Social conformity can be regarded as a choice situation in which the individual has to cope with conflicting demands Reliance upon his own opinions or perceptions, of course, means that the individual must risk social alienation He exposes him-

[7] These findings are part of a larger study of conforming behavior by Crowne and Liverant (1963)

self to the danger of standing alone and appearing different both to his peers and to the experimenter To conform, on the other hand, requires that he surrender his own beliefs to achieve alignment with the oppositional majority In Social Learning Theory the conformer can be characterized as one who has a low expectancy of success in socially evaluative situations, and he is thus inclined to use avoidant or defensive behavior in anticipation of disapproval and threat to his self-esteem In other words, conformity represents protective behavior designed to avert expected failure

The basic conformity situation for this experiment was similar to the procedure followed in the second experiment of Chapter 5 As before, the task set for the subjects required them to distinguish on each of 20 trials the larger of two groups of dots which were presented tachistoscopically for a short interval The naive subject and four preinstructed confederates of the experimenter were informed that the experiment was concerned with perceptual discrimination and speed On each trial they were required publicly to announce, in a given order, the letter designating the larger cluster of dots On 16 of the 20 trials the confederates unanimously announced incorrect judgments

The *M-C* scale was administered, as in some previous studies, on the first day of classes The measure of generalized expectancy was the Rotter (1942) Level of Aspiration Board This is a measure of goal-setting behavior and self-evaluation on which expectancies of future performance are compared with past successes and failures The subject is required to hit a little steel ball down a grooved board at the end of which is a series of numbers ranging from 1 to 10 and back down again to 1 The subject's task is to get as high a score as possible On each trial he must announce the score he expects to get Goal setting on this task may be described by a series of nine patterns (Rotter, 1954), each of which represents a characteristic approach to the setting of goals These patterns were our measure of generalized expectancy and were divided into two major groups The first of these groups, consisting of patterns 4 and 7, indicates over-cautiousness, self-protection, and defensive goal setting These patterns reflect lack of self-confidence and

avoidance of expected failure The second group of patterns, 1 and 3, represents a greater degree of confidence and more appropriate expectations of goal attainment In terms of the foregoing analysis, high-need-for-approval subjects with level-of-aspiration patterns 4 and 7 should constitute the conflict group and should display a much greater amount of conformity than low-need-for-approval subjects with level-of-aspiration patterns 1 or 3

Actually, this was a rather complex, multicondition experiment, and to achieve comparability among the experimental conditions, standard scores[8] computed separately for conditions and sexes were used to express the amount of conformity In standard scores, the mean amount of conformity of the high-need-for-approval/pattern 4, 7 group was 57 44 ($N = 9$) compared to a mean of 43 69 ($N = 13$) for the low-need-for-approval–pattern 1, 3 group The resulting t was 3 71, significant beyond the 01 level The mean difference here was much greater than that achieved by either the need-for-approval scale or the Level of Aspiration Board alone In fact, there was virtually no overlap between the two distributions, and prediction was equally successful for males and females

These results imply that the conformer can be regarded as an individual who has a high need for approval but a relatively low expectancy of success as a result of his own abilities and efforts His fear of social rejection results in a strong disposition to conform While there is no direct evidence bearing on the point, it is by no means clear that these subjects can explain the nature of the conflict to themselves and account for their own behavior What we seem to have accomplished in this experiment is a far more precise discrimination of avoidant, defensive, and threatened individuals—precisely those characteristics which emerge as descriptive of the approval-dependent person from the other vulnerable-self-esteem-and-defensiveness studies

There is quite a contrast between the failure-avoidant be-

[8] The mean of the standard score distribution was 50, and the standard deviation was 10

havior of approval-dependent individuals and their typical descriptions of themselves on personality measures Approval-motivated individuals, of course, tend to represent themselves in a highly positive and stereotypically acceptable manner, and the other findings reported in this section strongly suggest that they engage in defensive personal enhancement Indirectly supporting this line of reasoning are the findings of Barron (1953b) and Tuddenham (1959) that conformers tend to describe themselves in favorable terms on self-report devices

AFFILIATIVE FANTASY, EXPECTANCY OF REJECTION, AND UNPOPULARITY

One small study which appears to bridge the gap between dependence on the favorable evaluations of others and defensiveness and personal vulnerability was conducted at Dartmouth College [9] A group of 24 members of a small fraternity noted for its stress on group activities and sociability was recruited for a study in "Group dynamics" At the Dartmouth Human Relations Lab these subjects were given the *M-C* scale and five TAT cards (3BM, 6BM, 10, 4, and 6GF) The TAT was administered under standard nAffiliation instructions and later scored for affiliation thema and imagery according to the Heyns, Veroff, and Atkinson (Atkinson, 1958) manual Next, each subject was required to nominate in rank order five of his fraternity brothers as fitting each of two discriptions Description *A* included the following characteristics spends time with other people rather than alone, goes out of his way to make friends, very conversational, acts friendly Description *B* was spends much of his time alone rather than with other people, does not go out of his way to make friends, not very conversational, doesn't act friendly These paired descriptions were originally intended as a sociometric measure of overt affiliative behavior Postexperimental questioning, however, revealed that subjects actually made their ratings on the basis of popularity or like-

[9] One of us (Marlowe) directed the research project, which was conducted by Stephen P Bank

ability Thus, these nominations can best be considered as an index of the degree to which a fraternity member was positively regarded by his brothers This measure was scored by assigning a value of +5 for a 1-*A* nomination, +4 for a 2-*A*, etc, and by assigning a −5 for a 1-*B* nomination, −4 for a 2-*B*, etc No subject failed to receive nominations, and 20 of the 24 subjects received four or more

Correlations were then computed between the need-for-approval scale and the TAT measure of nAffiliation, and between the need for approval and the sociometric ratings of likeability A positive correlation of 55 was found between the need for approval and nAffiliation and a negative correlation of — 33 between the measure of approval dependence and the likeability ratings

The TAT stories were also scored for expectancy of rejection For each subject, those stories containing affiliation thema were examined for instances of obstacles or blocks in the path of affiliation and for evidences of negative affect These categories were defined according to the Heyns et al manual The expectancy-of-rejection score for each subject was the negative affect score plus the obstacle score for those stories containing affiliation thema This index was then divided by the number of stories containing affiliation thema to yield a score constant from subject to subject

Dividing the distribution of expectancy-of-rejection scores at the median and casting the data in a fourfold table (high versus low need for approval and above and below the median on expectancy of rejection), a Fisher Exact test yielded a probability of 10 The data of this analysis are shown in Table 4 High-need-for-approval subjects tended to give themes of rejection in their TAT stories to a greater degree than subjects with a lesser need for approval

A question arises in interpreting these findings whether the sociometric measure can be considered as a legitimate index of likeability Supporting the view that it was indeed a popularity or likeability measure is the significant negative correlation of — 42 between the nAffiliation measure and the sociometric index, a correlation closely replicating one obtained earlier by

TABLE 4 HIGH- AND LOW-NEED-FOR-APPROVAL SUBJECTS GIVING EXPECTANCY OF REJECTION THEMES

Expectancy of Rejection	Need for Approval	
	High	Low
High	7	2
Low	4	9

Fisher Exact test, $p = .10$.

Shipley and Veroff (1952) with a more clearly defined measure of likeability The likeability results have since been replicated by Bank, using a better sociometric measure of popularity-likeability He also found that high-need-for-approval subjects, to a significantly greater degree than lows, underestimated the number of "dislike" nominations they received

From these results, approval-motivated individuals appear to be affiliatively dependent, to anticipate social rejection in fantasy, and, in fact, to elicit unfavorable evaluations from others Why they are not liked is yet an open question, but others seem to react negatively to their encapsulated self-pictures, avoidant defenses, and dependence on external sources for reassurance and esteem The close associates of approval-seeking individuals appear to be able to recognize their defensiveness Barthel (1963) using the same nominating technique with defensive and nondefensive descriptions, found a tendency for high-need-for-approval individuals, especially those with a low generalized expectancy of success, to be rated by their fraternity brothers as more defensive in interpersonal relations. Approval-motivated individuals tended to rate themselves as less defensive than did low-need-for-approval subjects.

To recapitulate the findings of this triad of studies:

1 Approval-dependent patients tend to terminate psychotherapy earlier than less approval-dependent patients and without improvement. They are characterized as more defensive by

their therapists. Thus, their premature termination must be regarded as an avoidant and defensive maneuver

2 The conformity of approval-dependent persons appears to represent avoidant, self-protective behavior in anticipation of social rejection and threat to self-esteem

3 Despite their greater concern for affiliation with others as represented in a fantasy measure, subjects high in the need for approval tend to be less liked by peers who know them well and, moreover, tend to expect social rejection in fantasy, although not in response to direct questioning about their status in the group They are seen by their peers as defensive

11

Personality assessment: defensiveness, unrevealingness, and unproductivity on projective tests

We reviewed earlier, in Chapters 1 and 2, a portion of the abundant evidence that objective tests of personality are powerfully affected by test-irrelevant response determinants—so affected, in fact, as to make the assumption of an isomorphic relation between test responses and behavior in "real life" untenable. This realization dawned early to objective-test theorists and constructors, giving them claim to a certain distinction in recognizing the fallibilities of structured personality tests—and, of course, in laying the basis for the understanding of some of the factors underlying test responses.

When projective techniques came on the scene, the role accorded to them was to achieve a depth and accuracy of personality analysis to which structured instruments could not aspire—in L. K. Frank's (1939) bold analogy, a kind of psychic X ray baring the innermost structures and dynamics of personality. Such a view of projective tests conferred an immunity to both the conditions of test administration and the subject's test-taking attitudes and defenses. Neither could influence in any major way what the test would reveal. One effect of the projective or X-ray hypothesis was to stimulate a vast amount of re-

search seeking to confirm the consistency of projective measures of personality over an amazing variety of conditions and alterations of the state of the subject Masling (1960) wryly observed in his review of situational and interpersonal variables in projective testing that, "While there is no record of a Rorschach being given under water, there are reports of Rorschachs given under the influence of mescaline, alcohol, sodium amytal, hyoscine, morphine, and oxygen deprivation" Much of the early work on projectives, then, seems to have followed from the notion that motives in being tested, the testing context, particular instructions given to the subject, examiner characteristics, or altered physiological or psychological states could have little real effect in clouding the psychic X ray

Systematic research on projective techniques designed to assess the influence on projective responses of variables outside the subject's skin has been sketchy and guided by test theory to a far lesser degree than research on structured personality questionnaires Indeed, commitment to a particular view of projective tests is so intense that the nature and outcomes of projective-test studies themselves, notably investigations of the Rorschach, have been shown to be influenced by the setting in which the research was undertaken (Levy & Orr, 1959) Neither the avid proponents nor the ardent foes of projective tests have devoted themselves to a careful and objective analysis of the variables affecting projective tests

We can draw a rough parallel between research on the determinants of projective responses and the efforts to detect and to understand response bias and the conditions under which it occurs on personality questionnaires From an early and naively enthusiastic acceptance of projective-test responses as undistorted manifestations of unconscious dynamic processes, projective-test researchers undertook to determine the influence of faking, usually accomplished by means of instructions to create a good or bad impression (Carp & Shavzin, 1950, Fosberg, 1943, Weisskopf & Dieppa, 1951) These studies were followed by the introduction of situational variations stress (Lindzey, 1950a, 1950b, Lindzey & Herman, 1955), special instructions to the subject about the purpose of testing ("This is

a test to discover serious emotional disturbances," Henry & Rotter, 1956), and changes in the examiner's role (Luft, 1953) Then, with the advent of verbal conditioning, it was natural that attempts would be made to influence projective-test responses by the use of social reinforcers (Gross, 1959, Wickes, 1956)

Studies such as these have clearly established that projective methods are to a very marked degree affected by the goals and expectations of the subject in the test situation Projective responses may be changed by changing the meaning of the test situation or the importance of giving certain kinds of responses

In all of the research literature on projective tests scant attention has been paid to the effects of self-evaluative styles—individual differences in censorship or distortion—on minimally structured tests The impetus of the response-set research on personality inventories has led to a few studies essaying to determine the influence of a social-desirability response set on projective responses In one of the first of these investigations Rozynko (1959) found that psychiatric patients scoring high on a measure of social desirability tended to produce sentence completions having high social-desirability ratings He concluded that these patients felt threatened by the test and avoided revealing their own perceived weaknesses Another positive finding was reported by Reznikoff and Dollin (1961). In this study the TAT themes of a high-social-desirability group differed from those of low-social-desirability subjects in the preponderance of covert hostility over expressed aggression Thus, Reznikoff and Dollin's results appear to confirm the expected: the greater inhibition of hostility of subjects characterized by a tendency to respond in a socially desirable manner In a subsequent study, however, Reznikoff (1961) found no evidence of a social-desirability response set affecting TAT responses Similarly, Pena (1959) failed to find evidence of socially desirable responding on the Rorschach, and LeNoue, Spilka, Van de Castle, and Prince (1961) also concluded that the disposition to respond in a socially desirable manner is not carried over to the projective-test situation

Whether an individual's self-evaluative style as revealed in

responses to objective personality indices is also characteristically reflected on projective measures is still an open question The failure to show the effects of a social-desirability response set on projective tests is attributable to inadequacies of the measures on both the independent and dependent variable sides of the problem On the independent variable side, measures of the social-desirability response set differ from study to study They are simply not equivalent measures of the tendency to respond in a socially desirable way Further, many of them confound social desirability and the absence of psychopathology the subject is just not afflicted with the symptoms the items ask about Indices of the dependent variable in these studies —socially desirable projective responses—inadequately reflect avoidance, self-protection, and the tendency to respond according to some norm of social acceptability

Certainly the inconclusive findings of the response-set–projective-test studies to date cannot be taken as any kind of affirmation of the immunity of projective tests to the vicissitudes of censorship and distortion The problem lies, rather, in the failure to recognize that the ambiguity of projective tests is a double-edged sword It may be more or less successful in masking the nature of the test, but it also forces the subject to rely on minimal cues from the test instructions, the examiner, the context in which the testing is done, and the test stimuli in order to arrive at what is an appropriate response for him in this situation Rotter (1963) puts it this way

> Some studies suggest very strongly that the subject does not take an ambiguous test passively He attributes some purpose to the test, and the purpose he attributes to it may well hinge on odd bits of information, the setting of the test, differential emphasis on the words in the same instructions, and so forth In fact, it is just because the subject is not sure of the purpose of the test that his own hypotheses hinge on minor clues, frequently ones of which the examiner is not aware In short, as is strongly suggested by the study of Henry and Rotter [1956], these tests may be much more susceptible to slight differences in the conditions of testing than structured tests, and their utility may be sharply curtailed if the individuals using them are not aware of this susceptibility (pp 805–806)

The problem of social desirability in projectives is a complex one In addition to motivational differences among sub-

jects—differences in the *need* to respond in a socially desirable manner—there seem to be fewer salient cues in the projective-test situation as to *what is* socially desirable A lot of idiosyncratic responding, particularly in the content of responses, is the result In the studies reviewed above, social desirability on projectives has chiefly been indicated by judges' ratings of the content of percepts or themes

There appear to be, then, limited possibilities for agreement among subjects as to how to make the *content* of a response conform to a social-desirability criterion Projective cognoscenti, however, have long recognized the existence of defensive strategies Constriction and unproductivity, rigidity, naiveté and credulity, unrevealingness or barrenness, and stereotyped responding are readily identified as avoidant, self-protective approaches to being evaluated and are regarded as meaningful behavior in evaluating the nature of an individual's defenses (Schafer, 1948, 1954) If the disposition to respond in a socially desirable way is a widely generalized self-evaluative style, reflecting a dominant need for the approval of important others, we should expect it to be evidenced on projective tests just as in other evaluative situations How would such a defensive style be manifested? A clear expression of self-protection, avoidance of criticism, and a prevailing concern to defend one's self-image would be in constriction, limited productivity, and guardedness

In the remainder of this chapter we shall consider a study by Tutko (1962) designed to test this proposition Tutko predicted that the socially desirable self-evaluative style associated with the need for approval would result in constricted, defensive, unproductive, and ostensibly less pathological projective test protocols Two instructional conditions were used to test a second hypothesis concerning the effects of stress versus support or reassurance This prediction was that stressful instructions would lead to defensive responding more than would supportive test instructions Finally, Tutko predicted that the defensiveness of approval-dependent individuals would show up most clearly in the stressful instructional condition

Tutko's subjects were 60 psychotic or borderline psychotic

patients in a Veterans Administration hospital None of them were brain damaged, alcoholic, addicted to drugs, or retarded All of them had sufficient education to complete the tests, had been currently hospitalized less than five years, and were 45 years of age or less His subjects, then, were those functionally disturbed individuals for whom psychological test results would be important as a primary source of inference about personality dynamics and for the prediction of behavior This group was a singularly appropriate one on which to test his hypotheses

One of the two independent variables was the *M-C* scale The experimental conditions were created by means of two instructional sets, one stressful and the other supportive or reassuring Tested individually, each patient in the stressful condition was told "I am going to give you several psychological tests. It has often been found that the results of these tests can find out a great deal about people, and they are often used to find out if people are mentally sick Are you ready?" The supportive instructions assured anonymity and emphasized a concern with the tests themselves rather than an interest in the psychological scrutiny of the patient They concluded with the hopefully comforting statement "Whatever you say will have no bearing on your hospitalization in any way"

The projective tests were the Rorschach, TAT, and Incomplete Sentences Blank (ISB) (Rotter & Rafferty, 1950) The TAT was abbreviated and included those cards (1, 2, 4, 6BM, 7BM, 8BM, 15, 16, 13MF, and 18BM) most frequently employed in a clinical battery The standard clinical procedure was followed in administering the tests and, of course, Tutko counterbalanced the order of presentation

The projective-test protocols were coded, randomized, and then presented to 24 judges for rating[1] The 60 records of each test were divided in half Then, 30 protocols of one of

[1] Two groups of judges were used One group consisted of experienced clinicians and the other of less experienced graduate students in psychology In addition to the need-for-approval and instructional-set hypotheses, Tutko investigated the role of experience in making clinical judgments We will not be concerned with the analysis of this variable

the tests were given to each judge, to be rated on three scales Every test protocol was rated by four judges The rating task was to evaluate the degree of revealingness, pathology, and defensiveness of each test protocol These ratings were made on 9-point scales The revealingness scale can serve as an example The low point on this scale was defined as follows "This protocol will be considered extremely low in revealingness On the basis of the information given by the patient very little if any understanding of the patient's conflicts, emotional life, or personality structure can be determined This is a very 'restricted' record" A highly revealing test protocol (rating of 9) was defined as one in which "there will be few questions or doubts [about the patient's conflicts, emotional life, and personality structure] that could not be answered by the protocol This will be considered a 'rich' record" Similar types of definitions were given on the pathology and defensiveness scales

The subject's productivity on the three projective tests was also assessed On the Rorschach productivity was determined by counting the number of responses, and on the TAT and ISB the number of words per protocol was the productivity index Finally, Tutko hypothesized that human movement, color, and shading responses, which are considered to represent active fantasy, emotional responsiveness, and anxiety, would occur less frequently in the Rorschach records of approval-dependent patients A higher incidence of responses predicated on form and stereotyped and conventional content (animal and popular or common responses) was expected in the high-need-for-approval group

RESULTS

We first need to consider the reliability of the clinical ratings of the projective tests Interjudge reliability is summarized in Table 1, which shows the median intercorrelations found on each test and scale A fair degree of agreement was achieved on the revealingness and defensiveness scales, particularly when TAT or ISB protocols were being rated The Rorschach was the most difficult test to rate, and consensus on the degree of pathology

TABLE 1 Median Judge Intercorrelations on the Three Scales for Each Test

	Scale		
Test	Revealing-ness	Pathology	Defensive-ness
Rorschach	48 *	25	48 *
TAT	61 *	36 †	71 *
ISB	61 *	52 *	59 *

Note $N = 30$
$* p < 01$
$\dagger p < 05$

shown in test protocols proved elusive Generally, however, the clinicians were able to rate the tests on the three scales with a degree of reliability acceptable for the purposes of the study From the standpoint of reliability, the three scales, especially the revealingness and defensiveness scales, seem to be meaningful dimensions on which to evaluate projective test protocols

The analyses of the revealingness, pathology, and defensiveness scales were three four-way analyses of variance in which the independent variables were the experimental conditions, need for approval, tests, and experience[2] Table 2 shows the analysis of variance on the revealingness scale Significant main effects were found for both instructional set and the need for approval As predicted, the three tests tended to be rated as less revealing under the stress condition Nicely conforming to prediction were the findings with the need for approval Pa-

[2] Although there were some differences between experienced and inexperienced clinicians in the ratings assigned on certain of the scales and tests, these were in general relatively minor The pattern of the ratings given by the experienced and inexperienced judges was remarkably similar, and we shall do no violence to the findings to consider only the effects of the need for approval, instructional conditions, and type of test In the tables reporting the various analyses those portions dealing with experience are omitted

TABLE 2 Analysis of Variance for the Revealingness Scale

Source	*df*	*MS*	*F*	*p*
Instructional set	1	165 38	5 52	025
Need for approval	1	1480 28	49 38	001
Instructional set x need for approval	1	37 37	1 25	NS
Subjects within groups	56	29 98		
Tests	2	27 08	2 52	NS
Tests x instructional set	2	9 20	—	NS
Tests x need for approval	2	37 48	3 49	05
Tests x need for approval x instructional set	2	26 80	2 50	NS
Tests x subjects within groups	112	10 74		

tients high in the need for approval tended to produce projective-test protocols considerably less revealing than low-need-for-approval patients A significant interaction of tests and the need for approval was found, the results of which are depicted in Figure 1 In the high-need-for-approval group, the Rorschach was rated as significantly more revealing than either the TAT or ISB, which did not differ. Among low-need-for-approval patients, TAT protocols were rated as more revealing than ISB responses, but the TAT was not significantly different from the Rorschach With approval-dependent patients, then, the Rorschach was least susceptible to censorship, while the TAT and ISB were much more affected The projective protocols of low-need-for-approval patients were censored to a much lesser degree Contrasting the TAT and ISB protocols of this group, the TAT yielded material rated as much more personally revealing

On the analysis of variance of the pathology scale, which is presented in Table 3, a triple interaction was obtained Figure 2 shows this finding Neither the Rorschach nor the TAT yielded significant differences between the various groups (i e ,

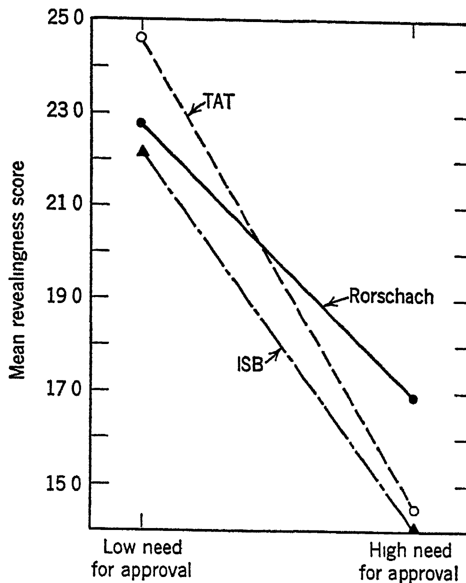


FIG 1 *Interaction on the revealingness scale tests by need for approval*

TABLE 3 ANALYSIS OF VARIANCE FOR THE PATHOLOGY SCALE

Source	*df*	*MS*	*F*	*p*
Instructional set	1	52 13	2 36	NS
Need for approval	1	3 02	—	NS
Instructional set x need for approval	1	56 81	2 57	NS
Subjects within groups	56	22 09		
Tests	2	36 96	5 00	05
Tests x instructional set	2	8 60	1 16	NS
Tests x need for approval	2	12 02	1 63	NS
Tests x need for approval x instructional set	2	35 01	4 74	05
Tests x subjects within groups	112	7.39		

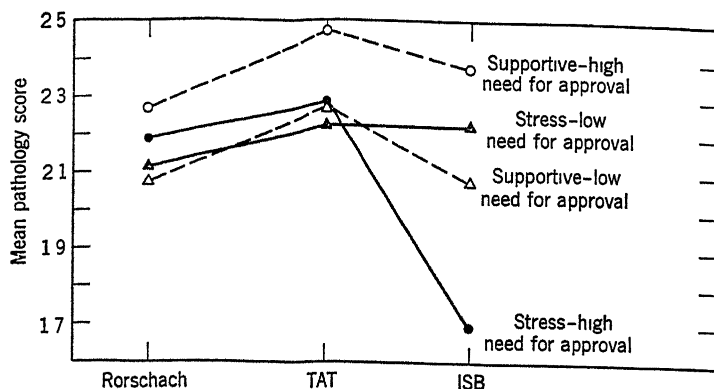


FIG 2 *Interaction on the pathology scale tests by need for approval by instructional set*

stress–high need for approval, supportive–low need for approval, etc) On the ISB, high-need-for-approval patients in the stress condition gave sentence completions rated significantly less pathological than those of any other group The ISB responses of low-need-for-approval patients in the supportive instructional condition were given significantly less pathological ratings than the completions of high-need-for-approval patients in the supportive condition Simply considering the amount of pathology present in a projective-test protocol, the ISB seems to be particularly influenced by the combined effects of the subject's motives in the test situation and instructions stressing the personal salience of the test results Under stress instructions, high-need-for-approval patients gave significantly less pathological sentence completions, suggesting the importance to the approval-dependent individual of presenting a defended, self-protective image of himself when personal evaluation is at stake Informing the subject that he is to be anonymous and is not to be personally evaluated appears to enable the approval-dependent patient to relax his defenses somewhat, and the extent of his disturbance is more clearly revealed

for-approval main effect, a significant interaction of the need for approval and experimental condition, and a test–by–need-for-approval interaction Conforming to expectation and consistent with the previous findings, the high-need-for-approval group was markedly more defensive than the low group The instructional-set–by–need-for-approval interaction is graphically presented in Figure 3 Among low-need-for-approval patients, the difference between stress and supportive instructions was not significant In the high group, the difference between instructional sets was significant beyond the 01 level With the more impersonal and anonymous instructions of the supportive condition, approval-dependent patients still gave responses rated as more defensive than patients less approval dependent This difference is dramatically increased in the stress condition As on the pathology ratings of the ISB, patients high in the need for approval were greatly affected by the stress instructions, becoming much more defensive

The final result, the interaction of tests with the need for approval, is depicted in Figure 4 In the high-need-for-approval group, TAT stories received significantly more defensive

TABLE 4 Analysis of Variance for the Defensiveness Scale

Source	*df*	*MS*	*F*	*p*
Instructional set	1	55 22	1 44	NS
Need for approval	1	1492 47	38 89	001
Instructional set x need for approval	1	153 41	4 00	05
Subjects within groups	56	38 38		
Tests	2	111 08	9 94	001
Tests x instructional set	2	8 64	—	NS
Tests x need for approval	2	40 48	3 62	05
Tests x need for approval x instructional set	2	26 18	2 34	NS
Tests x subjects within groups	112	11 17		

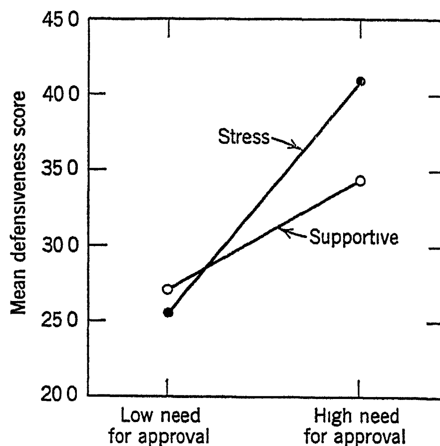


FIG 3 *Interaction on the defensiveness scale instructional set by need for approval*

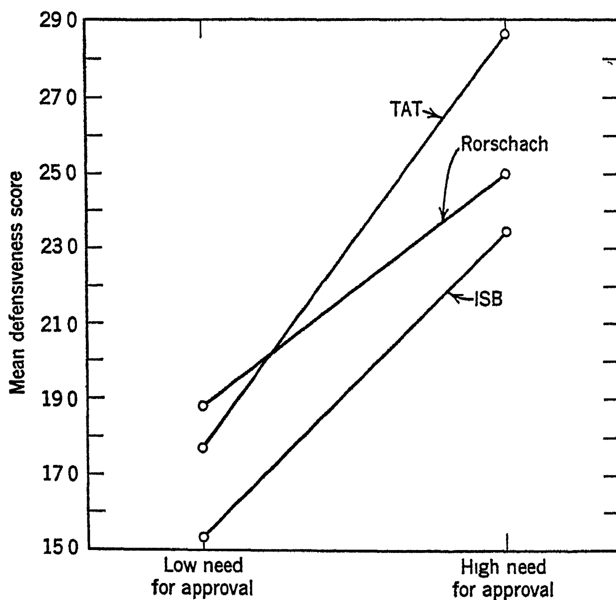


FIG. 4 *Interaction on the defensiveness scale tests by need for approval*

ratings than Rorschach or ISB responses, which did not differ In the low-need-for-approval group, the Rorschach protocols were significantly more defensive than ISB sentence completions, but the Rorschach did not differ from the TAT, nor was a significant difference found between the TAT and the ISB

Productivity was also assessed by means of analysis of variance The major finding was that high-need-for-approval patients were less productive on each test

Fantasy, color, shading, form-determined, animal, and popular responses on the Rorschach were scored as percentages of the total number of responses Since the first three were expected to be higher in the low-need-for-approval group and the last three higher among approval-dependent patients, the form, animal, and popular scores were subtracted from 100 to make the direction of scoring consistent An analysis of variance of these scores produced a significant main effect consistent with the prediction High-need-for-approval patients had lower scores—that is, they gave fewer Rorschach responses characterized by fantasy, color, or shading and more form-determined, animal, and popular responses On one of the two form measures, F+ (responses of good form quality as normatively determined), low-need-for-approval patients had higher scores Although giving more responses determined by form alone, high-need-for-approval patients tended to produce responses of poorer form quality Their responses, then, were more frequently those regarded as lacking in accuracy and good reality perception

As displayed in guardedness, conventional and stereotyped responses, and less productivity, the tendency to characterize oneself in a socially desirable manner appears clearly to be demonstrated on three of the most widely used projective tests Moreover, no test, even the unstructured and ambiguous Rorschach, is immune A subject does not abandon his goals and anticipations at the projective tester's door, nor, indeed, does he require a Baedeker to guide him through the intricacies of a projective test in order to portray himself, consciously or unconsciously, in a manner consistent with his own purposes A Baedeker helps, perhaps, as witnessed in the differential effect of the two instructional sets Stressing the personally evaluative consequences of

being tested resulted, for example, in greatly more defensive responding by approval-dependent patients. On the other hand, for the individual who is threatened by the prospect of being tested, instructions assuring him that he is not himself the focus of inquiry seem to result in somewhat less defensive responding.

How to avoid criticism, protect oneself, and defend one's self-image—in short, how to portray oneself in acceptable, if barren, terms on projective tests—are questions most readily answered by recourse to the behavior of high-need-for-approval subjects. (1) Give very few responses. (Two subjects in the high-need-for-approval group rejected ten and eight Rorschach cards, respectively. On the TAT, the uncommon occurrence of card rejection took place on two occasions, both times by high-need-for-approval patients.) (2) Respond to the obvious, stimulus-anchored characteristics of the test. (Consider the high incidence of popular responses and animal content on the Rorschach, the tendency for Rorschach responses to be determined by form, the paucity of fantasy and reflection, and the failure to respond to either the vivid or the variegated on the Rorschach blots.) (3) On the TAT, stories hewing close to simple card description and the avoidance of any imaginative themes make up the primary defensive strategy. (4) On the ISB, give innocuous completions.

It should not be concluded, however, that these are conscious or witting strategies employed by approval-motivated individuals. Although Tutko's data provide no evidence on the point, our earlier findings, particularly the results of the hostility-and-defense experiment reported in Chapter 9 and the findings of the psychotherapy-termination study in Chapter 10, suggest that the defensive style of the approval-dependent individual on projective tests, and perhaps on structured instruments, is not a thought-out and deliberately deceptive approach.

There are two perspectives in which the findings of this study may be regarded. One might take the traditional response-set position to argue that these findings reflect the amount of variance attributable to the tendency to respond in a socially desirable manner. Thus, the outcome is a projective-test protocol

less readily and meaningfully interpreted by virtue of unwanted and, unfortunately, unobviated contamination If, however, certain individuals bring the goals of self-protection and the avoidance of criticism to the test situation, and if, further, for such individuals a characteristic self-evaluative style acts in the service of those goals, what they do with their test responses may not be unmeaningful To recognize a defensive style and to develop inferences about its generality to other situations is an important and valuable endeavor for the projective tester Schafer (1954) cogently argues the point as follows

> In short, the patient not only wants to hide things from others but to hide them from himself as well In these desires he is no different from any of us, but, because typically he is in a life crisis, and because his overall integration is more or less precarious, he is likely to be especially threatened and defensive in a test situation How threatened he is by self-confrontation and how he copes with this threat are central aspects of his personality and pathology and will certainly be crucial factors in response to therapy (p 39)

The point probably need not be restricted to disturbed individuals and the prediction of therapeutic success, it might well serve to characterize the situation of most persons confronted with the prospect of being tested and the interpretive task facing the examiner

Defensive responding, then, is not simply a barrier in the interpretation of projective responses One approach to drawing meaningful inferences from projective-test responses of persons who display a socially desirable response set is represented by the recent effort to develop a manual to rate the need for approval itself from themes given in response to a modified TAT (Salman, 1964) The specific referents of need value for social approval include not only the individual's stylistic approach to being evaluated, but thematic content as well The following is a portion of Salman's manual for rating TAT stories

Need Value for Social Approval
Specific Referents

Strong Need for Social Approval

The foremost striving of the central character is to be loved and approved of

Need Value Rating

Central character is highly dependent on the favorable evaluation of others

Strong concern with the evaluative consequences of one's actions, i e, anticipation and concern about what other people might think, or with the impression one is making

The indiscriminate need to please others, to go along with the crowd and make favorable impressions in order to be thought well of and approved of

4 Moving toward people, a need to be with other people in order to gain their favor, find direction, or validate one's self-image in others

Reflected appraisals and the opinions of others have a marked effect in shaping the central character's self-concept, feelings of adequacy, etc Self-examination, reflection, or "soul searching" (strong personal involvement) following experienced rejection, being jilted, stood up, left out, or ignored (whereas indifference, lack of importance, or independence following such experiences will not be scored here)

Hero is unable to refuse doing favors, cannot say no, is unable to assert his authority because of his need for approval and a favorable evaluation

Central character is anxious not to offend anyone or hurt their feelings and therefore inhibits criticism, hostility, or self-assertion which might jeopardize approval satisfactions

Central character is forced to resort to unconventional, "devious," perhaps distasteful behavior in order to gain popularity, acceptance, or approval

Moderate Need for Social Approval

Marked concern with proper standards of behavior, the expression of conventional middle-class morality, and advocacy of a rigid code of conduct

Central character demonstrates marked concern with being "good," doing what is "right," and going "by the book" at all times

3 Characters in the story are judged or evaluated in the light of the social desirability of their actions, what others will think of them (e g, the neighbors)

Central character expresses contempt and derogates characters who are considered to be "improper" or "immoral"

Indeterminate, Doubtful, or Mild Suggestion of the Need for Approval

2 Defensive, highly unimaginative, and quite unrevealing stories

1 Mere inclusion of the word "approval" with no elaboration, the

remainder of the story questionable or unrelated to the need for approval

Independence of the Approval of Others

Central character's self-evaluation is not susceptible to, indeed may resist, manipulation by the reflected appraisals and opinions of other people

0 Greater concern with the intrinsic value of things rather than in their utility or instrumentality in obtaining social rewards, prestige, recognition, and a favorable evaluation

Negative attitudes, disapproval, or absence of consent from authority figures do not stand in the way of, or interfere with, the central character carrying out his plans

The following stories illustrate the application of some of the scoring criteria and suggest the kinds of dynamic content which may be found in the TAT themes of approval-motivated individuals The first story, told in response to a picture of two girls, one dressed in a raincoat with her hand on an open door and the other casually dressed in slacks with her back partially turned, is rated as a defensive, unimaginative, and unrevealing theme

There are two girls meeting in a corridor One girl is ready to go out and the other is coming in The girl going outside is all ready for rainy weather

One girl, in the rain attire, has read the newspaper or heard the weather predictions Since she is going to be outside quite a bit she prepares herself for bad weather

The girl in the rain attire is wondering just what is going on just as the girl without rain attire

The girl in the rain attire will peer outside and decide for herself whether or not there is the possibility of rain

The next story illustrates strong concern with the evaluative consequences of one's actions The picture is of a girl in an open doorway and a boy standing on the steps below her.

The boy has just presented the girl with a fraternity pin and she, not really liking him that much, is trying to decide what to do

She has led him on for social reasons Those being that she would like to be popular among boys

She is thinking that if she accepts the pin no other boy will want to ask her out But on the other hand the girls will be jealous

She will have a moment of guilt and decide not to accept the pin although she knows he will not ask her out again To her surprise he does ask her out again and she accepts because she realizes that she does like him and hopes he will ask her again

Note in this story the "moral," "socially desirable" decision of the heroine, which is rewarded The story ends on a happy note, unrealistically ignoring the earlier conflict Told in response to the same card, the following story also reflects an approval-dependence theme Here, reflected appraisals and the opinions of others strongly affect the hero's self-concept, his feelings of adequacy, and major decisions The powerful effects of the evaluations of others and the internalized importance of being good are suggested by the conclusion

The boy has just come to see his date as they had an argument the night before He gives her a present and she is embarrassed of his thoughtfulness The boy is apologizing but doesn't do much good

Last night when they went out, he started shouting at her for not being more sociable He took her to a party to influence his prospect of getting a job and she didn't make a favorable impression

He obviously realizes that he expected too much from her and in this helped to see his own faults He thinks that maybe he was wrong in taking the girl to make an impression—that he should be the one to do this

He will probably change his course in college and go into another field This situation will probably bring him to his senses

Contrast the stories above with the following theme, receiving a 0 score, in which the central character's self-evaluation is uninfluenced by others The frank expression of feelings of hostility is in clear evidence Family conflict is openly dealt with, and there is recognition of the motives and the frustrations of others The card to which this story is told is of an older man sitting on a couch holding a newspaper Standing behind him is a young girl of college age

Girl and her father Her pop is (was?) reading the paper with his feet up and his pipe near by

He had an exhausting day and when dinner was completely over decided to read the paper and go to bed

"Dad—could I have the front page for a current events idea—and Mom wants the society page and Tim the sports and the funnies too and Mother wants the garbage taken out and we'll be busy so you do it"

Contrary to the old story—Father will not be left paperless—he will say no—and on his way out knock over two lamps and a table

Although blandness, conventionality, and defensiveness may make the evaluation of projective material more difficult and although, as Tutko's data reveal, defensiveness may be increased in stressful testing situations, meaningful content may still be found Moreover, a major implication to be drawn from all of the defensiveness–and–vulnerable-self-esteem studies reported in Part III is that self-protection, the avoidance of criticism, and a strong concern to maintain a socially and personally acceptable image of oneself are manifested in a wide variety of interpersonal situations If this study further confirms that the projective test is something less than an X ray, it could not be said that the test *necessarily* fails to say something meaningful about the subject His defenses are meaningful, important, and widely generalized aspects of his personality, they are there to be seen in his approach to being evaluated

part IV Theory and integration

12

Theoretical issues and a broader analysis

The behavioral correlates of the tendency to avoid self-criticism and to choose self-evaluative statements which summatively portray a stereotypically acceptable self-image are complex, they occur in contexts far beyond the confines of the test situation. The presentation of a conforming public image in situations of self-evaluation has something of the quality of the masks worn by the characters in Eugene O'Neill's (1926) play, *The Great God Brown*—masks which the characters wear in self-protection and concealment of their inner selves from others. By the device of the masks, O'Neill sought to represent the imperative need to cover the nakedness of personal vulnerability from the insensitive forces of a world too ready to crush the defenseless. When the characters are introduced in the prologue of the play, they have already begun to learn, as youths, the lessons of dissimulation in the service of self-protection.

In our experimental analysis of the mask of social desirability, we found it necessary to scrutinize behavior in a variety of contexts less conspicuously involving the evaluation of self. Summarizing the results of the experiments described in the preceding chapters, individuals who display a social-desirability response set on the *M-C* scale are more conforming, cautious, and persuasible, and their behavior is more normatively anchored, than persons who depict themselves less euphemistically.

The greater amenability to social influence of persons who characterize themselves in very desirable terms is seen in (*a*) the favorability of their attitudes toward an extremely dull and boring task, (*b*) their greater verbal conditionability, both directly and vicariously, (*c*) social conformity, (*d*) a tendency to give popular word associations, (*e*) the cautious setting of goals in a risk-taking situation, (*f*) their greater reactivity, depending on their expectancies about the evaluative consequences of their behavior, in a "dirty word" perceptual-defense task, and (*g*) susceptibility to persuasion Moreover, among these individuals there is evidence to indicate a particular style of defense against hostility and self-protective, avoidant measures to avert anticipated threats to self-esteem The latter eventuates in a "leaving the field" form of resistance to psychotherapy Finally, approval-dependent persons seek affiliation but tend to be disliked Awareness of how they are evaluated seems to emerge only on a fantasy measure

The findings thus suggest that more is involved than a contrived and deliberate presentation of self in the sense of Goffman's (1959) analysis of "impression management" in social relationships They imply that self-evaluation, dependence on the favorable esteem of others, and lack of assertiveness act in the service of self-protection, avoidance, and defense Taken together, the studies of the approval motive suggest a set of self-reflexive attitudes—a self-conception—in which an idealized version of the self (Horney, 1950) is maintained and defended Many of the behaviors which are associated with defensive self-evaluation appear to follow from, to be closely articulated with, and to support the approval-dependent person's self-esteem

In presenting this interpretation, we have taken sides in a dialectic going back to the origins of personality testing, in which are pitted the opposing notions of conscious and deliberate faking or misrepresentation and a less frankly aware, defensive kind of self-depiction The studies on the approval motive are not, of course, definitive at this point, but they do suggest the latter

SOME THEORETICAL ISSUES

The experimental analysis of the approval motive conforms to the familiar R-R paradigm (Bergmann & Spence, 1944) in the formal establishment of relationships between one set of responses (self-characterization on a test) and another diverse array of responses At the simplest level, what these findings have demonstrated is a relation between a self-evaluative style and "other-directed" (Riesman, 1950), self-protective, and conflict-avoidant behavior The formulation of the need-for-approval construct and its corollaries of defensiveness and vulnerable self-esteem was intended to give meaning and theoretical significance to responses on the *M-C* scale and to bridge the gap between test responses and nontest behavior With each of the dependent variables investigated, the relationship was specified in a hypothesis derived from the concept, and each experiment was undertaken to test a theoretically based prediction

This is the logic and strategy of construct validity (Cronbach & Meehl, 1955) the attempt to give meaning to test responses and to converge on a particular, defined, and discriminant interpretation In the validation of a construct, a single relationship can give the inference of a motive or dispositional state no more than heuristic support, but with the enlargement of a nomological net embodying a set of relationships with diverse criteria, a stronger claim is established In a sense, the essential problem in construct validity, as in any theoretical endeavor, is to rule out alternative interpretations The process goes by various names "convergent operationalism" (Garner, Hake, & Eriksen, 1956), "methodological triangulation" (Campbell, 1956), and "convergent validation" (Campbell & Fiske, 1959) The core of the problem to which these terms refer is that a single set of experimental operations can often be interpreted in many different ways. To converge on one explanatory concept—to establish that a relationship is explained within a particular nomological net and cannot be so adequately accounted for in another net—requires a process of "triangula-

tion" from a number of different operations In construct validity, the more extensive the domain of empirical relationships, given that they are specified by and consistent within the theory, the closer is the approximation to convergence on a single and unambiguous interpretation

We can see the problem clearly in trying to interpret the relationship between the *M-C* scale and attitudes toward the uninteresting, dull, and personally meaningless spool-packing task This relationship we interpreted as validating evidence for the approval motive Edwards, Diers, and Walker (1962), however, have contended that the *M-C* scale is better conceived as a measure of lying In citing these results, they have argued that our high-scoring subjects were actually prevaricating when they characterized the task so favorably Choosing between these competing interpretations of the meaning of the scale is perhaps largely a matter of one's preference But when other relationships are added, the feasibility of such an alternative interpretation may be more seriously questioned It would be difficult to hold, for example, that the verbal-conditioning experiments confirm the Edwards et al "liar" hypothesis, there is really no basis on which to predict that liars would be any more or less influenced by subtle social reinforcement than those more committed to the truth The verbal-conditioning experiments are much better explained by the need for approval, which also explains the implied demand study In the conjunction of these two sets of findings, a step has been taken toward the elimination of competing alternatives

The problem of increasing convergence on an explanation is complicated by the fact that in the prediction of very similar behaviors even a large number of relationships may not provide the breadth of empirical support necessary to rule out competing possibilities We saw examples of this in reviewing some of the studies of test-taking behavior Here, although the amount of research has been enormous, the variables investigated have been almost entirely restricted to self- and item evaluation The test literature burgeons with examples of alternative and in most cases equally plausible accounts of the response-set problem in personality inventories In large part

the difficulties encountered are attributable to the restricted number of experimental or assessment operations that have been employed. Operations have paralleled each other on the independent- and dependent-variable sides of the problem; investigators have failed to attain conceptual "distance" between independent and dependent variables. There is a prevailing similarity between the sets of responses investigated.

One of the major issues in construct validity lies in the theoretical link between a construct and the test presumed to reflect it. In defining the relation between an individual's responses to personality-test items and the significance to be attached to his responses it has been conventional to make one of two assumptions. These assumptions concern the importance of item content.

First, one may assume that the content of items is irrelevant. What is important, then, is simply whether or not an item is capable of discriminating empirically between criterion and normative groups. Whatever the item may mean to the respondent is of no concern as long as the predictive purpose is fulfilled. This, of course, is the "dust bowl" empirical approach of the MMPI. If one makes this assumption about the importance or meaning of item content in personality-test construction, no further assumptions are necessary.

The second assumption is that subjects respond to the *content* of personality-test items and evaluate themselves with reference to the item material presented to them. Thus, agreement or disagreement with a personality-test item is mirrored in the individual in certain attributes, traits, or habitual ways of responding. A crude analogy from psychophysics can illustrate the point. It is as if each item constitutes a standard stimulus against which the subject is to compare himself with respect to the trait or characteristic in question. To be sure, there may be other determinants of test responses, such as a tendency to endorse favorable items and to reject those which would project an unfavorable image. However, subjects are still assumed to be affected primarily by the content of the items, and various devices (matching forced-choice items for social desirability) may be employed to obviate other influences. Edwards' Per-

sonal Preference Schedule clearly relies on this second assumption, and the Taylor Manifest Anxiety scale may be cited as a second case in point

The *M-C* scale rests on neither of these assumptions In distinction to the second assumption, there is no assumed isomorphism between item content and what is characteristic of the individual Thus, individual differences in the need for approval are inferred from self-characterization but not, as we earlier pointed out, by assuming that the subject is describing his behavior accurately Rather, favorable self-evaluation on the *M-C* scale is assumed to reflect an individual's habitual response style and the goals and expectations which are aroused in situations of self-evaluation The scale, then, is assumed to measure a person's approach to self- and socially evaluative situations and the meaning that such situations have for him Item content is not altogether irrelevant In fact, it is necessary that test items involve the kind of content and tone or style which will engage those goals and expectations pertinent to being evaluated Items may vary widely as long as this characteristic is maintained

The *M-C* scale sets up an evaluative problem for the subject, and we infer his need for approval from the way he handles it There is an analogy here to the linkage between test responses and the inferences derived from them in projective testing The inference of achievement motivation (or other motive states) from imaginative stories proceeds from the projective hypothesis that the subject gives his own meanings to the test stimuli and to the situation The themes that appear in a subject's stories testify to his major concerns, preoccupations, and desires Unknown to himself, the hypothesis goes, he gives himself away, betraying his needs by the way he handles this evaluative problem We have chosen to make the *M-C* scale more frankly self-evaluative and more highly structured than is typically the case with the stimuli used in projective testing The core of the similarity lies in the stylistic analysis of the individual's approach to the test and the inferential steps between his responses and the motives which promote their choice

To summarize, the *M-C* scale supposes that a person brings

to the test situation a habitual pattern or style of evaluating himself and that items appropriate to self-evaluation will tend to call forth responses reflecting his particular style From his approach to the test situation we may infer a closely interwoven motivational structure centering around dependence on the favorable evaluations of others and a vulnerable self-conception

The final and crucial step in triangulation will be to anchor the concept of the approval motive in terms of antecedent conditions—to determine the childhood experiences giving rise to defensive self-evaluation, conformity, and the other behaviors mediated by the need for approval Not only is this an essential step in the validation of the approval motive, in addition, we stand to extend our knowledge of child development This is, however, a difficult undertaking Research on the childhood antecedents of personality variables faces some formidable obstacles In studying the development of dispositional states three basic procedures can be followed, each of which poses its own problems

First, the research strategy pursued in the studies reported in this volume might be extended downward in investigations of approval dependence in children Thus, a children's version of the *M-C* scale might be developed to predict behavior on dependent variables such as those we have employed with college students Although this cross-sectional procedure has merit, it does entail the assumption that a need-for-approval scale for children has the same meaning for them that it has for adults A further complexity might lie in the nonlinear development of approval dependence and its corollary dynamics from early childhood to adulthood We raise here the intriguing possibility that dramatic changes may take place in the development of the motive and that the defensive structure which appears to be an important part of the approval motive may not appear until somewhat later in the course of socialization

A second procedure in the investigation of the childhood antecedents of the approval motive would be to obtain retrospective accounts of childhood experiences and parental behavior, attitudes, and child-rearing practices from adults (e g , college stu-

dents or psychiatric patients) and to attempt to distinguish between the recollections of high- and low-need-for-approval individuals This process, of course, closely resembles the clinical method of studying personality development There is, however, a critical difficulty to which the concept of the approval motive is especially vulnerable We could expect the reconstructions of childhood by high-need-for-approval individuals to be biased by the suppression—or, indeed, repression—of negative attitudes toward parents Thus, from the defensive and avoidant dynamics of the approval-dependent person we would predict attempts to avert conflict by the idealization of parents and childhood A very distorted picture would be the result

The third approach involves a longitudinal research program The difficulties of longitudinal personality research are obvious: obtaining a sample on which repeated measures can be taken, maintaining contact with subjects, both parents and children, over a long period of time; the choice of reliable and theoretically meaningful variables, and, finally, an interim before the final outcome can be known that is subjectively interminable to many investigators of personality.

These problems are not raised to demean research on childhood antecedents nor to suggest that the obstacles are insuperable The issues are compelling, however, and they ought to be realistically joined before studies are undertaken The research strategy in the work on the approval motive has been based on the assumption that a meaningful network of contemporaneous findings ought to be established first. A theoretically articulated set of findings will better direct the search for antecedents A further concern is to establish the viability of the measure of approval dependence and the hypotheses linking it to behavioral dependent variables before attempting to accomplish the major step in triangulation

Where one draws the line and decides that a network of R-R relationships is sufficiently compelling to warrant the investigation of childhood determinants is perhaps largely a matter of one's impetuosity or cautiousness The predictive power of the need for approval does seem to be sufficiently well established,

however, to make it feasible for future studies to attempt the systematic investigation of its antecedent conditions.

THE APPROVAL MOTIVE AND SOCIAL LEARNING THEORY

The remaining task is an analysis of the approval motive in the conceptual language of Rotter's (1954) Social Learning Theory The logic of this attempt requires a bit of justification It is obvious that the research on the need for approval stands in its own right, both in the empirical demonstration of a series of relationships and the meaning given to them The need for approval has the status of a miniature theory devised to account for a small fraction of the domain of human behavior, and it may be accepted or dismissed depending on its ability to predict and to organize an interrelated set of findings

There are, nonetheless, some clear advantages to incorporating the construct in a more inclusive personality theory First of all, a personality theory can yield both a breadth of view and possibilities of internal articulation of concepts from which new and unanticipated findings may emerge The work on nAchievement and other motives (Atkinson, 1958, McClelland, Atkinson, Clark, & Lowell, 1953) has surely profited enormously from the development of a general theory of motivation Both the generalizability of the approach to other dispositional states than achievement and the increased theoretical refinement of expectancy and incentive concepts can be seen as consequences of a more general appproach In the same way the Social Learning Theory analysis of the approval motive led to strikingly increased accuracy in the prediction of conforming behavior in the experiment reported in Chapter 10

Second, the incorporation of a construct within the larger ground of a personality theory may help to avoid a problem to which research in personality is especially vulnerable This is the difficulty created by a cluttered taxonomy of motives, in which motivational constructions are offered for each new facet of behavior whose scrutiny is undertaken A systematic personality theory can offer the advantage of parsimony in speci-

fying only a limited and reasonable number of dispositional variables

Allied to this is the reverse problem of imperialism—attempting to account for everything with a single variable, thus failing to recognize the dangers involved in the dilution of a concept and the necessity to achieve discriminant validity (Campbell & Fiske, 1959) In a critical review of research on self-acceptance, Crowne and Stephens (1961) tried to show the problems encountered in theoretical imperialism which grow out of the failure to specify the limiting parameters of a construct In this problem too, the perspective of a personality theory may be advantageous A general behavior theory is less likely to emphasize the all-inclusive character of a given theoretical concept, and it is probable that a systematically explicated theory will make provision for a number of concepts

These considerations, then, suggested the desirability of taking a stand on some basic issues in the psychology of motivation and argued for an attempt to place the concept of approval dependence within the realm of a more general theory

Basic Concepts in Social Learning Theory

Social Learning Theory has been extensively described elsewhere (Rotter, 1954, 1955), and its implications in several areas of human behavior have been explored as well (Liverant, 1963, Rotter, 1960, 1962, Rotter, Seeman, & Liverant, 1962) For our purposes, it will suffice to outline the basic concepts of the theory and their interrelations

In the tradition of Tolman (1932) and Lewin (1935) and concordant with the earlier ideas of McDougall (1923), Social Learning Theory emphasizes that behavior is purposive, striving, and goal-directed The theory rejects the consummatory determinants of drive-reduction theories It is much more in keeping with the cognitive trend of recent psychological theory summarized in the writings of White (1959, 1960). Arguing that drive reduction provides an insufficient base for the explanation of complex, goal-oriented behavior, White (1960) seeks to emphasize

> manipulation, locomotion, language, the building of cognitive maps and skilled actions, and the growth of effective behavior in relation to other people These acquisitions are made by young animals and children partly through exploratory and manipulative play when drives such as hunger and sex are in abeyance The directed persistence of such behavior warrants the assumption of a motivation independent of drives, here called effectance motivation, which has its immediate satisfaction in a feeling of efficacy and its adaptive significance in the growth of competence Effectance motivation can be likened to independent ego energies in the psychoanalytic scheme The child's actual competence and his sense of competence are built up from his history of efficacies and inefficacies, and a sense of competence is held to be a crucial element in any psychology of the ego By directing attention to action and its consequences and to the vicissitudes of the sense of competence, it should help to speed the construction of an adequate ego psychology (pp 137–138)

Rotter's theory has much the same focus, and the concept of freedom of movement, as we shall see, is concerned with phenomena very similar to those of interest to White

In Social Learning Theory, behavior is a function of learned goals and subjective probabilities regarding their attainment which are acquired in the course of social interaction If the initial learning of the human infant is influenced by physiologically based drives, learned motives soon take precedence and function autonomously to direct subsequent behavior The prediction of complex social behavior can be accounted for without postulating states of drive or processes of tension reduction In Rotter's terms, "We find it a sufficient basis for prediction to state that behavior directed toward the attainment of a learned goal or external reinforcement may be predicted through a knowledge of the situation the organism is in and from a knowledge of his past learning experiences" (1954, p 116)

The basic concepts of the theory specify that the probability of occurrence of a given behavior (Behavior Potential) is a joint function of the expectancy that certain reinforcing events will result from the behavior and the value of, or preference for, the reinforcements The schematic equation is, accordingly, $BP = f(E \& RV)$ It is important to note that expectancy does not refer to objective probability but to the individual's

subjective anticipations about the likelihood of particular outcomes based upon his past experiences. The concept of reinforcement value is an empirical one and refers to the acquired or learned values, positive or negative, of the outcomes or consequences of behavior.

The basic equation above refers to the prediction of a single behavior as a result of a particular expectancy and reinforcement value. In the prediction of those broader categories of behavior of interest in the study of personality, three new terms are introduced. A set of behaviors may be functionally related in their orientation toward a given goal—a set of related reinforcements. The tendency of a person to behave in a particular way in a number of different situations can be taken as an indication that his behavior is directed toward the achievement or avoidance of certain reinforcements which these situations share in common. In Social Learning Theory, a functionally related set of behaviors directed toward the attainment of a goal defines a need. The probability or potentiality of occurrence of such a related set of behaviors is expressed by the concept of need potential. As in the basic formula above, need potential is determined by the combination of expectancy and reinforcement value, here termed freedom of movement and need value. Freedom of movement refers to the generalized expectancy or probability of success in the attainment of a functionally related set of reinforcements, and need value is the relative preference for those reinforcements. Schematically, NP = f(FM & NV).

The psychological situation, as we saw in Chapter 2, plays an important role in providing cues for an individual's expectancies regarding the reinforcements to be attained by alternative behaviors. In the course of psychological development, the individual learns to discriminate situations that involve different reinforcements or reward different behaviors, and he also learns to generalize from one situation to another when the reinforcements or reinforced behaviors are similar. Social Learning Theory takes account of both situationally specific expectancies and generalized or trans-situational expectancies such as a per-

son's learned ways of evaluating himself and his feelings of adequacy or inadequacy in many social contexts

Definition of a Motive

As implied in the foregoing discussion, Social Learning Theory defines a motive or need behaviorally as the probability of occurrence of certain behaviors directed toward the achievement of a goal As observed above, it is not necessary to assume the reduction of drives or to postulate, as do McClelland et al (1953), a primary affective state and processes of learning and redintegration in order to account for the directionality of behavior Neither does Rotter make use of a concept of *psychological* tension involving displacements from, restorations of, and the general tendency of psychological processes toward equilibrium, as Lewin (1936) did

How many needs are there in the conceptual scheme which Social Learning Theory presents? The determination of major goals or needs is an empirical problem, and a conceptualization of particular motives or needs must ultimately follow from the demonstration of functional relatedness among sets of reinforcements Rotter, however, has suggested six broad need areas which would appear to provide a preliminary description of significant goals These are needs for recognition-status, protection-dependency, dominance, independence, love and affection, and physical comfort Preliminary research by Rockwell (1950) and the development by Liverant (1958) of an inventory measuring the relative importance of certain of these need areas suggest the promise of this provisional description of psychological needs Our conception of approval dependence does not fit exactly into this scheme, and it appears to cut across several of the need areas The behaviors associated with the need for approval are complex and multi-faceted, and it might be expected that a number of different goals would be involved —goals which are functionally related through the individual's self-evaluation and the meaning that his relationships with other persons have for his conception of himself

Incorporation

The approval motive, as conceived in Social Learning Theory terms, is defined by the concept of need potential Our analysis makes clear that both generalized expectancies (freedom of movement) and need value are involved in dependence on the favorable evaluations of others and in avoidance of self-criticism. From the concert of needs and generalized expectancies derives the desire for social support, self-protection, and avoidance of failure that we have labeled the approval motive

The goals or needs of the approval-dependent person would appear to include social recognition and status, protection and dependency, and love and affection However, there has been no systematic attempt to determine the more specific goals involved in the approval motive, and these suggested need values can only be regarded as tentative The sociometric study of likeability reported in Chapter 10 suggests that affiliation, or love and affection in the terminology of Social Learning Theory, is a singularly important goal in the constellation of needs comprising the approval motive

It seems reasonable to assume that the high-need-for-approval individual has learned that conformity, submission, and the normative anchoring of his behavior entail the fewest risks of social rejection and threats to self-esteem His self-justification and attempts to validate his own self-worth seem to represent defensive efforts to cope with anticipated failures We may assume, thus, low freedom of movement for independence, assertiveness, and the frank acknowledgement of feelings The sensitivity to social rejection of approval-dependent individuals is suggested by the prevalence of expectancy-of-rejection themes found in their TAT stories in the likeability study We have some evidence, therefore, to support the hypothesis of low freedom of movement. Individuals dependent upon the favorable evaluations of others appear to have difficulty in recognizing and contending with hostility As earlier suggested, it may be that their unrecognized anger and resentment lead to their being less liked by others The low freedom of movement of approval-de-

pendent persons, then, appears to result in failure-avoidant and defensive behavior in situations involving a conflict between needs to be approved and accepted on the one hand and both internalized and perceived social prohibitions on the other

The conformity study reported in Chapter 10 attempted to identify defensive and failure-avoidant level-of-aspiration patterns among approval-motivated subjects in an effort to reveal the effects of low freedom of movement more clearly In this experiment, social conformity was regarded as behavior designed to avert the conflict between, on one hand, the desire to be right and, on the other, lack of confidence in one's abilities and the perceived risks of criticism and social isolation entailed by the independent assertion of one's own beliefs Such a conflict would be especially intense for the person dependent on others for approval, respect, and affection whose evaluation of his ability to achieve his goals is low This pattern—a high need value for certain reinforcements and low expectancies of their attainment—represents the Social Learning Theory approach to maladjustment In this sense, maladjusted behavior is self-protective, serving to avert expected failure directly or symbolically We saw that those approval-dependent subjects characterized as having a high expectancy of failure tended to resolve the conflict by submission to the group The behavior of these individuals was in compelling contrast to a group less dependent on approval and with higher expectancies of success

Another example of self-protective resolution of conflict was seen in the therapy study, where high-need-for-approval patients tended to resolve the avoidance-avoidance conflict posed by the perceived necessity of defying the therapist or meeting the requirement of self-disclosure by leaving the field It appears that they expected that both these alternative behaviors would lead to unpleasant consequences The avoidance of hostility demonstrated by the Conn study is yet an additional example of defensive conflict resolution among approval-motivated individuals in which high expectancies of punishment for aggression may be assumed Finally, we may mention the prediction of rigid and defensive goal setting in Barthel's (1963) experiment Here, as in the conformity study, the addition of a meas-

ure of generalized expectancy of success resulted in the more accurate prediction of rigid behavior

The research on the approval motive and the Social Learning Theory conceptualization of the findings cast the problem of maladjustment in an interesting light In a sense, high-need-for-approval individuals are more "normal" in that they exemplify many of the values of American middle-class culture Approval-motivated individuals say the right things about themselves, appear to hold "proper" attitudes, reflect common language usage in their associations, set goals of acceptably intermediate risk, do not show hostility, and seem in general to reflect the congeries of virtues defining the "adjusted" individual On closer scrutiny, however, those identified as approval dependent seem to resolve some social and personal conflicts in ways that result in detriment to themselves As traditionally conceived, maladjustment is associated with deviance from cultural norms—personal dissatisfaction and self-rejection, acknowledgement of uncommon symptoms, and inappropriate social behavior (Scott, 1958) Our findings might be taken to suggest that those who are in many respects least deviant by the usual criteria of adjustment may be, in fact, rather maladjusted individuals The "autoplastic adaptations" of high-need-for-approval individuals, although leading to conformity to social norms, are, in the conceptual terms outlined here, maladjustive orientations involving the avoidance of failure and punishment which appear to result in the unsatisfactory solution of problems and the perpetuation of conflict Thus, our results and theory suggest that the criteria of personal unhappiness, acknowledgement of symptoms, and deviant social behavior are fallible in the definition of maladjustment More to the point, perhaps, would be to recognize the importance of the individual's goals and his expectancies of success or failure in achieving them

Summing up, research needs to be devoted to the separate measurement of the need and freedom-of-movement components of the approval motive, both to enhance prediction of behavior and to gain a better understanding of the dynamics involved Procedures for determining expectancies of success similar to those used in the conformity study might be followed, or new

devices might be developed to assess both need value and freedom of movement The thematic-apperception measure described in the last chapter may be a promising technique for the assessment of need value and freedom of movement

The Social Learning Theory analysis of the approval motive is largely a heuristic one at this juncture, but it raises a host of intriguing possibilities for subsequent exploration The separate measurement of need value and freedom of movement should help to determine the degree and extent of the defensiveness of approval-motivated persons and may bear additional fruit in clarifying the intrapersonal processes involved in such complex social behavior as conformity and persuasibility

13

Epilogue

We started, in the beginning of this book, with the problem of response sets on personality inventories and in particular with the problem of a social-desirability response set. The uncertainty, the lack of clarity, the feculence surrounding the question of the meaning of personality-test responses led us to abandon the ground which seemed clear to others—and so muddy to us—and to strike out more or less on our own. Thus, the internal analysis of tests was forsaken; instead, we directed our search toward the goals and expectations that would impel one to evaluate himself in terms conditioned by the acceptance of others. To do so required us to postulate a motivational state, reflected in test-taking behavior, and to seek its correlates in behaviors less harassed by the confusions of personality tests.

Our findings have been confirmative, although in the process a major alteration of the concept of the approval motive—the defensiveness–and–vulnerable-self-esteem hypothesis—was necessary to account for some unanticipated and initially paradoxical results. Testing further some of the implications of the new hypothesis led to positive findings. The association of defensiveness and protection of self-esteem with dependence on the approval of others appears sufficiently well established now to warrant its formal acceptance as an integral part of the construct.

In a sense, then, this book traces the natural history of a

miniature theory from its limited and timorous beginnings through a program of research entailing a significant reconceptualization We have tried to convey to the reader a sense of our own discovery by describing our findings very nearly in the order in which they happened With only a few exceptions, we have presented our research as we have done it and our thinking virtually as we have thought it Scientific inquiry has more than a little of the heady anticipation, excitement, and mystery of a good detective yarn—if one can stay conceptually abreast of his data Such has been our good fortune so far

But a story ends, while research does not It is true perforce that such an Odyssey is never complete The experiments described in the foregoing chapters only sketchily outline the meanings of the approval motive There is a great deal to be done filling in the gaps left in the domain we have traversed, testing more extended implications of the construct, and anchoring it on the antecedent-conditions side of the problem The normative-anchoring findings need further research Conventionality in language and dependence on external sanction in social attitudes and behavior should be given closer scrutiny to determine the reliability and generality of these findings For example, can we predict culturally stereotyped language habits in everyday speech from our knowledge of the tendency of approval-dependent individuals to give common word associations? The defensiveness-and-vulnerable-self-esteem hypothesis should be tested further, as we suggested in the last chapter, using the double classification of need value and freedom of movement of Social Learning Theory

Appendix: Marlowe-Crowne social-desirability scale norms

Below we report normative data on the *M-C* scale for a number of the populations that we have studied. Detailed norms in the form of standard scores and centile equivalents are given for The Ohio State University introductory psychology students, since this is the population from which we have drawn the largest proportion of our experimental samples. For other groups the number of cases, mean, and standard deviation are reported. Because there is a consistent though slight difference between males and females on each sample, the norms are given separately.

Normative data are reported for several samples not investigated in the studies presented in this volume. These data were collected by other psychologists who very kindly provided them to us, or were gathered by us subsequent to the studies in this book. They are presented here in the hope of stimulating investigations of these and other populations with the scale.

THE OHIO STATE UNIVERSITY

These were introductory psychology students tested on the first day of classes or at the time of participating in an experiment. There is no difference in score between these methods of testing

	Males				Females		
	N = 666				N = 752		
	Mean = 15 06				Mean = 16 82		
	SD = 5 58				SD = 5 50		
Raw Score	Number of Cases	Centile Equivalent	Standard Score*	Raw Score	Number of Cases	Centile Equivalent	Standard Score*
0	1	0	23	0	0	—	19 †
1	3	1	25	1	0	—	21 †
2	1	1	27	2	1	0	23
3	4	1	28	3	2	0	25
4	2	2	30	4	2	1	27
5	13	4	32	5	4	1	29
6	13	6	34	6	7	2	30
7	22	9	36	7	15	4	32
8	16	11	37	8	17	6	33
9	33	16	39	9	21	9	36
10	37	22	41	10	31	13	38
11	38	27	43	11	33	18	39
12	44	34	45	12	40	23	41
13	35	39	46	13	42	29	43
14	53	47	48	14	46	35	45
15	44	54	50	15	48	41	47
16	48	61	52	16	63	49	49
17	33	66	53	17	52	56	50
18	47	73	55	18	37	61	52
19	38	79	57	19	50	68	54
20	27	83	59	20	39	73	56
21	22	86	61	21	43	79	58
22	26	90	62	22	44	85	59
23	15	92	64	23	30	89	61
24	17	95	66	24	23	92	63
25	14	97	68	25	17	94	65
26	4	98	70	26	12	96	67
27	6	98	71	27	12	97	69
28	9	99	73	28	7	98	70
29	1	99	75	29	6	99	72
30	0	—	77 †	30	4	99	74
31	0	—	79 †	31	3	99	76
32	0	—	80 †	32	1	99	78
33	0	—	82 †	33	0	—	79 †

* Mean of 50, standard deviation of 10 † Extrapolated

Sample	Sex	Number of Cases	Mean	*SD*
Northwestern University				
Introductory psychology students	Male	100	11 65	5 26
	Female	86	13 51	4.75
Harvard University				
(a) Paid volunteers for psychological experiments	Male	45	13 02	3 46
(b) Graduate students	Male	19	14 35	5 61
Dartmouth College				
Introductory psychology students	Male	32	10 06	4.37
University of Washington				
Undergraduate students	Male	110	14 39	5 62
University of North Dakota				
Volunteers for psychological experiments	Male	49	13 88	4 95
	Female	59	16 04	4 44
Lesley College				
Freshmen and Sophomores	Female	60	14 20	4 62
Boston University				
Undergraduate fraternity members	Male	41	13 73	4 63
Secretarial school students	Female	60	16 27	5 53
Insurance Company				
(a) Employed women (told that scores would be seen only by experimenter and not by anyone connected with company)	Female	88	15 42	6 16
(b) Applicants (led to believe that scores would be considered in hiring)	Female	285	24 62	4 96

Sample	Sex	Number of Cases	Mean	*SD*
California prisoners	Male	80	16 73	6 04
Massachusetts prisoners				
(a) Professional prostitutes	Female	17	21 41	12 78
(b) Professional and amateur prostitutes combined	Female	26	19 11	11 39
(c) Definitely not prostitutes	Female	24	16 33	6 38
Psychiatric inpatients in Veterans Administration hospital (diagnosed schizophrenic)	Male	60	16 48	6 65
Psychiatric clinic outpatients (both neurotic and psychotic diagnoses)	Male	40	12 20	6 44
	Female	46	11 54	6 59

References

Adorno, T W , Frenkel-Brunswik, Else, Levinson, D J , & Sanford, R N 1950 *The authoritarian personality* New York Harper

Allison, J , & Hunt, D E 1959 Social desirability and the expression of aggression under varying conditions of frustration *J consult Psychol*, **23**, 528–532

Allport, G W 1928 A test for ascendance-submission *J abnorm soc Psychol*, **23**, 118–136

Allport, G W , & Vernon, P E 1933 *Studies in expressive movement* New York Macmillan

Asch, S E 1951 Effects of group pressure upon the modification and distortion of judgments In H Guetzkow (Ed), *Groups, leadership and men, research in human relations* Pittsburgh Carnegie Press

Asch, S E 1952 *Social psychology* New York Prentice-Hall

Atkinson, J W (Ed) 1958 *Motives in fantasy, action, and society* Princeton, N J Van Nostrand

Atkinson, J W , Bastian, J R , Earl, R W , & Litwin, G H 1960 The achievement motive, goal setting, and probability preferences *J abnorm soc Psychol*, **60**, 27–36

Atkinson, J W , Heyns, R W , & Veroff, J 1954 The effect of experimental arousal of the affiliation motive on thematic apperception *J abnorm soc Psychol*, **49**, 405–410

Atkinson, J W , & Litwin, G H 1960 Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure *J abnorm soc Psychol*, **60**, 52–63

Atkinson, J W , & Reitman, W R 1956 Performance as a function of motive strength and expectancy of goal-attainment *J abnorm soc Psychol*, **53**, 361–366

Ax, A F 1953 The physiological differentiation between fear and anger in humans *Psychosom Med*, **15**, 433–442

Barron, F 1953a An ego-strength scale which predicts response to psychotherapy *J consult Psychol*, **17**, 327–333
Barron, F 1953b Some personality correlates of independence of judgment *J Pers*, **21**, 287–297
Barthel, C E 1961 The need for approval and goal-setting behavior Unpublished research, The Ohio State University
Barthel, C E 1963 The effects of the approval motive, generalized expectancy, and situational cues upon goal setting and social defensiveness Unpublished doctoral dissertation, The Ohio State University
Barthel, C E, & Crowne, D P 1962 The need for approval, task categorization, and perceptual defense *J consult Psychol*, **26**, 547–555
Bass, B M. 1955 Authoritarianism or acquiescence? *J abnorm soc Psychol*, **51**, 616–623
Bass, B M, & Berg, I A (Eds) 1959 *Objective approaches to personality assessment* Princeton, N J · Van Nostrand
Berg, I A 1955 Response bias and personality The deviation hypothesis *J Psychol*, **40**, 61–72
Berger, S M 1961 Incidental learning through vicarious reinforcement *Psychol Rep*, **9**, 477–491
Bergmann, G, & Spence, K W 1944 The logic of psychophysical measurement *Psychol Rev*, **51**, 1–24
Birney, R C 1958 Thematic content and the cue characteristics of pictures In J W Atkinson (Ed), *Motives in fantasy, action, and society* Princeton, N J Van Nostrand
Bitterman, M E, & Kniffin, C W 1953 Manifest anxiety and "perceptual defense" *J abnorm soc Psychol*, **48**, 248–252
Blake, R R, & Brehm, J W 1954 The use of tape recording to stimulate a group atmosphere *J abnorm soc Psychol*, **49**, 311–313
Buss, A H 1961 *The psychology of aggression* New York Wiley
Cady, V M 1923 The estimation of juvenile incorrigibility *J delinqu Monogr*, No 2.
Campbell, D T 1956 Leadership and its effects upon the group Monogr No 83 Columbus The Ohio State Univer Bur Business Res.
Campbell, D T, & Fiske, D W 1959 Convergent and discriminant validation by the multitrait-multimethod matrix *Psychol Bull*, **56**, 81–105.
Carp, A L, & Shavzin, A R 1950 The susceptibility to falsification of the Rorschach diagnostic technique *J consult Psychol*, **14**, 230–233
Chapman, L J, & Campbell, D T 1957. Response set in the F scale *J abnorm. soc Psychol*, **54**, 129–132
Clark, R A 1952 The projective measurement of experimentally induced levels of sexual motivation *J exp Psychol*, **44**, 391–399
Cofer, C N, Chance, June, & Judson, A J 1949 A study of malingering on the MMPI *J Psychol*, **27**, 491–499

Cohen, A R 1959 Some implications of self-esteem for social influence In C I Hovland & I L Janis (Eds), *Personality and persuasibility* New Haven Yale Univer Press

Conn, L K, & Crowne, D P 1964 Instigation to aggression, emotional arousal, and defensive emulation *J Pers*, **32**, 163–179

Couch, A, & Keniston, K 1960 Yeasayers and naysayers Agreeing response set as a personality variable *J abnorm soc Psychol*, **60**, 151–174

Couch, A, & Keniston, K 1961 Agreeing response set and social desirability *J abnorm soc Psychol*, **62**, 175–179

Cowen, E L, & Tongas, P N 1959 The social desirability of trait descriptive terms Applications to a self-concept inventory *J consult Psychol*, **23**, 361–365

Crandall, V J 1951 Induced frustration and punishment-reward expectancy in thematic apperception stories *J consult Psychol*, **15**, 400–404

Cronbach, L J 1946 Response sets and test validity *Educ psychol Measmt*, **6**, 475–494

Cronbach, L J 1950 Further evidence on response sets and test design *Educ psychol Measmt*, **10**, 3–31

Cronbach, L J, & Meehl, P E 1955 Construct validity in psychological tests *Psychol Bull*, **52**, 281–302

Crowne, D P 1959 The relation of self-acceptance behavior to the social learning theory construct of need value Unpublished doctoral dissertation, Purdue University

Crowne, D P, & Liverant, S 1963 Conformity under varying conditions of personal commitment *J abnorm soc Psychol*, **66**, 547–555

Crowne, D P, & Marlowe, D 1960 A new scale of social desirability independent of psychopathology *J consult. Psychol*, **24**, 349–354

Crowne, D P, & Stephens, M W 1961 Self-acceptance and self-evaluative behavior A critique of methodology *Psychol Bull*, **58**, 104–121

Crowne, D P, & Strickland, Bonnie R 1961 The conditioning of verbal behavior as a function of the need for social approval *J abnorm soc Psychol*, **63**, 395–401

Crutchfield, R S 1955 Conformity and character *Amer Psychologist*, **10**, 191–198

Dollard, J, & Miller, N E 1950 *Personality and psychotherapy* New York McGraw-Hill

Dollard, J, & Mowrer, O H 1947 A method of measuring tension in written documents *J abnorm soc Psychol*, **42**, 3–32

Dunn, Sandra, Bliss, Joan, & Siipola, Elsa 1958 Effects of impulsivity, introversion, and individual values upon association under free conditions *J Pers*, **26**, 61–76

Edwards, A L 1953a The relationship between the judged desirability of a trait and the probability that the trait will be endorsed *J appl Psychol*, **37**, 90–93

Edwards, A L 1953b *Manual for the Edwards Personal Preference Schedule* New York Psychol Corp

Edwards, A L 1957 *The social desirability variable in personality assessment and research* New York Dryden

Edwards, A L, & Diers, Carol J 1962 Social desirability and conflict *J soc Psychol*, **58**, 349–356

Edwards, A L, Diers, Carol J, & Walker, J N 1962 Response sets and factor loadings on sixty-one personality scales *J appl Psychol*, **46**, 220–225

Edwards, A L, & Walker, J N 1961a A note on the Couch and Keniston measure of agreement response set *J abnorm soc Psychol*, **62**, 173–174

Edwards, A L, & Walker, J N 1961b A short form of the MMPI The SD scale *Psychol Rep*, **8**, 485–486

Edwards, A L, & Walker, J N 1961c Social desirability and agreement on response set *J abnorm soc Psychol*, **62**, 180–183

Ellis, A 1946 The validity of personality questionnaires *Psychol Bull*, **43**, 385–440

Eriksen, C W 1951 Perceptual defense as a function of unacceptable needs *J abnorm soc Psychol*, **46**, 557–564

Festinger, L 1954 A theory of social comparison processes *Hum Relat*, **7**, 117–140

Festinger, L 1957 *A theory of cognitive dissonance* Evanston, Ill Row, Peterson

Festinger, L, & Carlsmith, J M 1959 Cognitive consequences of forced compliance *J abnorm soc Psychol*, **58**, 203–210

Fosberg, I A 1943 How do subjects attempt fake results on the Rorschach test? *Rorschach Res Exch*, **7**, 119–121

Frank, L K 1939 Projective methods for the study of personality *J Psychol*, **8**, 389–413

Freud, S 1949 Further recommendations in the technique of psychoanalysis On beginning the treatment The question of the first communications The dynamics of the cure In *Collected papers*, Vol II London Hogarth Press Originally published in 1913

Fujita, B 1956 An investigation of the applicability of the PPS to a cultural subgroup Unpublished master's thesis, University of Washington

Galton, F 1879 Psychometric experiments *Brain*, **2**, 149–162

Garner, W R, Hake, H W, & Eriksen, C W 1956 Operationism and the concept of perception *Psychol Rev*, **63**, 149–159

Goffman, E 1959 *The presentation of self in everyday life* Garden City, New York Doubleday

Gough, H G 1948 A new dimension of status I Development of a personality scale *Amer sociol Rev*, **13**, 401–409

Gough, H G 1951 Studies of social intolerance II A personality scale for antisemitism *J soc Psychol*, **33**, 247–255

Grant, D A 1956 Analysis-of-variance tests in the analysis and comparison of curves *Psychol Bull*, **53**, 141–154

Greenspoon, J 1955 The reinforcing effect of two spoken sounds on the frequency of two responses *Amer J Psychol*, **68**, 409–416

Gross, L R 1959 Effects of verbal and nonverbal reinforcement in the Rorschach *J consult Psychol*, **23**, 66–68

Haber, R N, & Alpert, R 1958 The role of situation and picture cues in projective measurement of the achievement motive In J W Atkinson (Ed), *Motives in fantasy, action, and society* Princeton, N J Van Nostrand

Hartshorne, H, & May, M A 1928 *Studies in the nature of character* Vol I *Studies in deceit* New York Macmillan

Heider, F 1958 *The psychology of interpersonal relations* New York Wiley

Henry, Edith, & Rotter, J B 1956 Situational influence on Rorschach responses *J consult Psychol*, **20**, 457–462

Hollingshead, A B, & Redlich, F C 1958 *Social class and mental illness* New York Wiley

Hoppe, F 1930 Etfolg und misserfolg *Psychol Forsch*, **14**, 1–62

Horney, Karen 1950 *Neurosis and human growth* New York Norton

Horton, D L, Marlowe, D, & Crowne, D P 1963 The effect of instructional set and need for social approval on commonality of word association responses *J abnorm soc Psychol*, **66**, 67–72

Hovland, C I, & Janis, I L (Eds) 1959 *Personality and persuasibility* New Haven Yale Univer Press

Hovland, C I, Janis, I L, & Kelley, H H 1953 *Communication and persuasion* New Haven Yale Univer Press

Howes, D H, & Solomon, R L 1950 A note on McGinnies' "Emotionality and perceptual defense" *Psychol Rev*, **57**, 229–234

Humm, D G, & Humm, Kathryn A 1944 Validity of the Humm-Wadsworth temperament scale With consideration of the effects of subjects' response-bias *J Psychol*, **18**, 55–64

Humm, D G, Storment, R C, & Iorns, M E 1939 Combination scores for the Humm-Wadsworth temperament scale *J Psychol*, **7**, 227–253

Humm, D G, & Wadsworth, G W 1935 The Humm-Wadsworth temperament scale *Amer J Psychiat*, **92**, 163–200

Jackson, D N, & Messick, S J 1958 Content and style in personality assessment *Psychol Bull*, **55**, 243–252

Jackson, D N, & Messick, S J 1961 Acquiescence and desirability as response determinants on the MMPI *Educ psychol Measmt*, **21**, 771–790

Jackson, D N, Messick, S J, & Solley, C M 1957 How "rigid" is the "authoritarian"? *J abnorm. soc Psychol*, **54**, 137–140

Janis, I L 1954 Personality correlates of susceptibility to persuasion *J Pers*, **22,** 504–518

Janis, I L, & King, B T 1954 The influence of role-playing on opinion change *J abnorm soc Psychol*, **49,** 211–218

Jenkins, J J 1959 Effects on word-association of the set to give popular responses *Psychol Rep*, **5,** 94

Jenkins, J J 1960 Commonality of association as an indicator of more general patterns of verbal behavior In Sebeok, T A (Ed), *Style in Language* Cambridge, Mass Technology Press

Jung, C G, & Riklin, F 1904 Diagnostische Assoziation studien I Experimentelle Untersuchungen uber Assoziationen gesunder *J Psychol Neurol*, **3,** 193–215

Kagan, J 1956 The measurement of overt aggression from fantasy *J abnorm soc Psychol*, **52,** 390–393

Kanfer, F H, & Marston, A R 1963 Human reinforcement vicarious and direct *J exp Psychol*, **65,** 292–296

King, B T, & Janis, I L 1956 Comparison of the effectiveness of improvised versus non-improvised role playing in production of opinion changers *Hum Relat*, **9,** 177–186

Klett, C J 1956 A study of the Edwards Personal Preference Schedule in relation to socio-economic status Unpublished doctoral dissertation, University of Washington

Krasner, L 1958 Studies of the conditioning of verbal behavior *Psychol Bull*, **55,** 148–170

Krasner, L, Ullmann, L P, Weiss, R L, & Collins, Beverly J 1960 A personality correlate of individual differences in responsivity to verbal conditioning as obtained by three examiners Paper read at Western Psychol Assoc, San Jose, Calif

Lacey, O W, Lewinger, N, & Adamson, J 1953 Foreknowledge as a factor affecting perceptual defense *J exp Psychol*, **45,** 169–174

LeNoue, D, Spilka, S, Van deCastle, R, & Prince, A 1961 Social desirability and the group Rorschach *J clin Psychol*, **17,** 175–177

Leventhal, H, & Perloe, S I 1962 A relationship between self-esteem and persuasibility *J abnorm soc Psychol*, **64,** 385–388

Levy, L H, & Orr, T B 1959 The social psychology of Rorschach validity research *J abnorm soc Psychol*, **58,** 79–83

Lewin, K 1935 *A dynamic theory of personality* New York McGraw-Hill

Lewin, K 1936 *Principles of topological psychology* New York McGraw-Hill

Lewis, D J, & Duncan, C P 1958 Vicarious experience and partial reinforcement *J abnorm soc Psychol*, **57,** 321–326

Lindzey, G 1950a An experimental examination of the scapegoat theory of prejudice *J abnorm soc Psychol*, **45,** 296–309

Lindzey, G 1950b An experimental test of the validity of the Rosenzweig Picture Frustration Study *J Pers*, **18**, 315–320

Lindzey, G, & Herman, P S 1955 Thematic apperception test A note on reliability and situational validity *J proj Tech*, **19**, 36–42

Linton, Harriet, & Graham, Elaine 1959 Personality correlates of persuasibility In C I Hovland & I L Janis (Eds), *Personality and persuasibility* New Haven Yale Univer Press

Liverant, S 1958 The use of Rotter's social learning theory in developing a personality inventory *Psychol Monogr*, **72**, No 2 (Whole No 455)

Liverant, S 1963 Learning theory and clinical psychology In L E Abt & B F Riess (Eds), *Progress in clinical psychology* Vol V New York Grune & Stratton

Liverant, S, & Scodel, A 1960 Internal and external control as determinants of decision making under conditions of risk *Psychol Rep*, **7**, 59–67

Luft, J 1953 Interaction and projection *J proj Tech*, **17**, 489–492

McClelland, D C, Atkinson, J W, Clark, R A, & Lowell, E L 1953 *The achievement motive* New York Appleton-Century-Crofts

McDougall, W 1923 *Outline of psychology* New York Scribner

McGee, R K 1962 Response style as a personality variable By what criterion? *Psychol Bull*, **59**, 284–295

McGinnies, E 1949 Emotionality and perceptual defense *Psychol Rev*, **56**, 224–251

McKinley, J C, Hathaway, S R, & Meehl, P E 1948 The MMPI VI The K scale *J consult Psychol*, **12**, 20–31

Madison, P 1961 *Freud's concept of repression and defense* Minneapolis Univer of Minnesota Press

Mangan, G L, Quartermain, D, & Vaughn, G M 1960 Taylor MAS and group conformity pressure *J abnorm soc Psychol*, **61**, 146–147

Marlowe, D 1959 Some personality and behavioral correlates of conformity Unpublished doctoral dissertation, The Ohio State University

Marlowe, D 1962 Need for social approval and the operant conditioning of meaningful verbal behavior *J consult Psychol*, **26**, 79–83

Marlowe, D, Beecher, R S, Cook, J B, & Doob, A N 1964 The approval motive, vicarious reinforcement and verbal conditioning Paper read at Eastern Psychol Assoc, Philadelphia, Penna

Marlowe, D, & Crowne, D P 1961 Social desirability and response to perceived situational demands *J consult Psychol*, **25**, 109–115

Marlowe, D, Stifler, L, & Davis, M 1962 Personality correlates of primary and secondary suggestibility Unpublished research, Dartmouth College

Marston, W M 1920 Reaction time symptoms of deception *J exp Psychol*, **3**, 72–87

Masling, J 1960 The influence of situational and interpersonal variables in projective testing *Psychol Bull*, **57**, 65–85

Meehl, P E 1945 The dynamics of "structured" personality tests *J clin Psychol*, **1**, 296–303

Meehl, P E, & Hathaway, S R 1946 The K factor as a suppressor variable in the MMPI *J appl Psychol*, **30**, 525–564

Merrill, R M, & Heathers, Louise B 1956 The relation of the MMPI to the Edwards Personal Preference Schedule on a college counseling center sample *J consult Psychol*, **20**, 310–314

Messick, S, & Jackson, D N 1961 Acquiescence and the factorial interpretation of the MMPI *Psychol Bull*, **58**, 299–304

Miller, N E, & Dollard, J 1941 *Social learning and imitation* New Haven Yale Univer Press

Mussen, P H, & Naylor, H K 1954 The relationships between overt and fantasy aggression *J abnorm soc Psychol*, **49**, 235–240

Mussen, P H, & Scodel, A 1955 The effect of sexual stimulation under varying conditions on TAT sexual responsiveness *J consult Psychol*, **19**, 90

Newcomb, T M 1959 Individual systems of orientation In S Koch (Ed), *Psychology A study of a science* Vol III *Formulations of the person and the social context* New York McGraw-Hill

Norman, W T 1963 Relative importance of test item content *J consult Psychol*, **27**, 166–174

O'Neill, E 1926 *The great God Brown, The fountain, The moon of the Caribbees, and other plays* New York Boni and Liveright

Orne, M T 1959 The nature of hypnosis Artifact and essence *J abnorm soc Psychol*, **58**, 277–299

Orne, M T 1962 On the social psychology of the psychological experiment with particular reference to demand characteristics and their implications *Amer Psychologist*, **17**, 776–783

Osgood, C E, & Tannenbaum, P H 1955 The principle of congruity in the prediction of attitude change *Psychol Rev*, **62**, 42–55

Pena, C D 1959 Influence of social desirability upon Rorschach content *J clin Psychol*, **15**, 313–316

Postman, L, Bronson, Wanda C, and Gropper, G L 1953 Is there a mechanism of perceptual defense? *J abnorm soc Psychol*, **48**, 215–224

Reznikoff, M 1961 Social desirability in TAT themes *J proj Tech*, **25**, 87–89

Reznikoff, M, & Dollin, Adelaide 1961 Social desirability and the type of hostility expressed on the TAT *J clin Psychol*, **17**, 315–317

Riecken, H 1962 A program for research on experiments in social psychology In N F Washburne (Ed), *Decisions, values, and groups* Vol II London Pergamon Press

Riesman, D 1950 *The lonely crowd* New Haven Yale Univer Press

Rockwell, Anne F 1950 The evaluation of six social learning need constructs Unpublished doctoral dissertation, The Ohio State University

Rogers, J M 1960 Operant conditioning in a quasi-therapy setting *J abnorm soc Psychol*, **60**, 247–252

Rosen, E 1956 Self-appraisal, personal desirability, and perceived social desirability of personality traits *J abnorm soc Psychol*, **52**, 151–158

Rosenberg, M J 1960 An analysis of affective-cognitive consistency In M J Rosenberg and C I Hovland (Eds), *Attitude organization and change* New Haven Yale Univer Press

Rosenberg, M J 1961 Proposal for continuation of a research program on consistency and change in social attitudes Unpublished manuscript, The Ohio State University

Rosenthal, R 1963 On the social psychology of the psychological experiment the experimenter's hypothesis as unintended determinant of experimental results *Amer Scientist*, **51**, 268–283

Rosenzweig, S 1934 A suggestion for making verbal personality tests more valid *Psychol Rev*, **41**, 400–401

Rotter, J B 1942 Level of aspiration as a method of studying personality II Development and evaluation of a controlled method *J exp Psychol*, **31**, 410–422

Rotter, J B 1943 Level of aspiration as a method of studying personality III Group validity studies *Charact & Pers*, **11**, 254–274

Rotter, J B 1954 *Social learning and clinical psychology* New York Prentice-Hall

Rotter, J B 1955 The role of the psychological situation in determining the direction of human behavior In M R Jones (Ed), *Nebraska symposium on motivation* Lincoln, Nebr Univer of Nebraska Press

Rotter, J B 1960 Some implications of a social learning theory for the prediction of goal directed behavior from testing procedures *Psychol Rev*, **67**, 301–316

Rotter, J B 1962 Some implications of a social learning theory for the practice of psychotherapy Unpublished manuscript, The Ohio State University

Rotter, J B 1963 A historical and theoretical analysis of some broad trends in clinical psychology In S Koch (Ed), *Psychology A study of a science* Vol V *The process areas, the person, and some applied fields Their place in psychology and in science* New York, McGraw-Hill

Rotter, J B, & Rafferty, Janet E 1950 *Manual for the Rotter Incomplete Sentences Blank* New York Psychological Corp

Rotter, J B, Seeman, M R, & Liverant, S 1962 Internal versus external control of reinforcement A major variable in behavior theory In N F Washburne (Ed), *Decisions, values, and groups* Vol II London Pergamon Press

Rozynko, V V 1959 Social desirability in the sentence completion test *J consult Psychol*, **23**, 280

Ruch, F L 1942 A technique for detecting attempts to fake performance on a self-inventory type of personality test In Q McNemar and Maude A Merrill, *Studies in personality* New York McGraw-Hill

Salman, A R 1962 The need for approval, improvisation, and attitude change Unpublished master's thesis, The Ohio State University

Salman, A R 1964 The development and validation of a projective measure of need for approval Doctoral dissertation in progress The Ohio State University

Salzinger, K , & Pisoni, Stephanie 1960 Reinforcement of verbal affect responses of normal subjects during the interview *J abnorm soc Psychol ,* **60,** 127–130

Sarason, I G 1958 Interrelations among individual difference variables, behavior in psychotherapy, and verbal conditioning *J abnorm soc Psychol ,* **56,** 339–344

Sarason, I G 1959 Relationships of measures of anxiety and experimental instructions to word association test performance *J abnorm soc Psychol ,* **59,** 37–42

Schachter, S , & Singer, J E 1962 Cognitive, social, and physiological determinants of emotional state *Psychol Rev ,* **69,** 379–399

Schachter, S , & Wheeler, L 1962 Epinephrine, chlorpromazine, and amusement *J abnorm soc Psychol ,* **65,** 121–128

Schafer, R 1948 *The clinical application of psychological tests* New York International Univer Press

Schafer, R 1954 *Psychoanalytic interpretation in Rorschach testing* New York Grune & Stratton

Scott, W A 1958 Research definitions of mental health and mental illness *Psychol Bull ,* **55,** 29–45

Sears, Pauline S 1940 Levels of aspiration in academically successful and unsuccessful children *J abnorm soc Psychol ,* **35,** 498–536

Sears, R R , Maccoby, Eleanor E , & Levin, H 1957 *Patterns of child rearing* Evanston, Ill Row, Peterson

Seeman, J 1954 Counselor judgments of therapeutic process and outcome In C R Rogers & Rosalind F Dymond (Eds), *Psychotherapy and personality change* Chicago Univer of Chicago Press

Shipley, T E , & Veroff, J 1952 A projective measure of need for affiliation *J exp Psychol ,* **43,** 349–356

Siipola, Elsa, Walker, W Nannette, & Kolb, Dorothy 1955 Task attitudes in word association, projective and non-projective *J Pers ,* **23,** 441–459

Spielberger, C D , Berger, A , & Howard, Kay 1963 Conditioning of verbal behavior as a function of awareness, need for social approval, and motivation to receive reinforcement *J abnorm soc Psychol ,* **67,** 241–246

Spielberger, C D , Levin, S M , & Shepard, Mary 1962 The effects of

awareness and attitude toward the reinforcement on the operant conditioning of verbal behavior J Pers, **30,** 106–121

Strickland, Bonnie R 1962 The relationship of awareness to verbal conditioning and extinction Unpublished doctoral dissertation, The Ohio State University

Strickland, Bonnie R, & Crowne, D P 1962 Conformity under conditions of simulated group pressure as a function of the need for social approval *J soc Psychol*, **58,** 171–181

Strickland, Bonnie R, & Crowne, D P 1963 Need for approval and the premature termination of psychotherapy *J consult Psychol*, **27,** 95–101

Taffel, C 1955 Anxiety and the conditioning of verbal behavior *J abnorm soc Psychol*, **51,** 496–501

Taylor, Janet A 1953 A personality scale of manifest anxiety *J abnorm soc Psychol*, **48,** 285–290

Taylor, J B 1961 The "yeasayer" and social desirability A comment on the Couch and Keniston paper *J abnorm soc Psychol*, **62,** 172

Tolman, E C 1932 *Purposive behavior in animals and man* New York Appleton-Century

Tuddenham, R D 1959 Correlates of yielding to a distorted group norm *J Pers*, **27,** 272–284

Tutko, T A 1962 Need for social approval and its effect on responses to projective tests Unpublished doctoral dissertation, Northwestern University

Ullmann, L P, Krasner, L, & Collins, Beverly J 1961 Modification of behavior through verbal conditioning Effects in group therapy *J abnorm soc Psychol*, **62,** 128–132

Weiner, D N 1948 Subtle and obvious keys for the Minnesota Multiphasic Personality Inventory *J consult Psychol*, **12,** 164–170

Weiss, R L, Ullmann, L P, & Krasner, L 1960 On the relationship between hypnotizability and response to verbal operant conditioning *Psychol Rep*, **6,** 59–60

Weisskopf, Edith, & Dieppa, J J 1951 Experimentally induced faking of TAT responses *J consult Psychol*, **15,** 469–474

Welsh, G S 1956 Factor dimensions A and R In G S Welsh & W G Dahlstrom (Eds), *Basic readings on the MMPI in psychology and medicine* Minneapolis Univer of Minnesota Press

Wheeler, W M, Little, K B, & Lehner, G F J 1951 The internal structure of the MMPI *J consult Psychol*, **15,** 134–141

White, R W 1959 Motivation reconsidered The concept of competence *Psychol Rev*, **66,** 297–333

White, R W 1960 Competence and the psychosexual stages of development In M R Jones (Ed), *Nebraska symposium on motivation* Lincoln, Nebr Univer of Nebraska Press

Whyte, W H 1956 *The organization man* New York Simon & Schuster

Wickes, T A 1956 Examiner influence in a test situation *J consult Psychol*, **20,** 23–26

Wiggins, J S 1959 Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions *J consult Psychol*, **23,** 419–427

Wiggins, J S 1962 Strategic, method, and stylistic variance in the MMPI *Psychol Bull*, **59,** 224–242

Wiggins, J S, & Rumrill, C 1959 Social desirability in the MMPI and Welsh's factor scales A and R *J consult Psychol*, **23,** 100–106

Wolf, S, & Wolff, H G 1947 *Human gastric function* New York Oxford Univer Press

Author index

Subject index